

MODELS OF CLASSICAL AND RECURRENT NOVAE

M. Friedjung and H. Duerbeck

I. INTRODUCTION.

I.A. BASIC THINGS TO BE EXPLAINED

The behavior of novae may be divided roughly into two separate stages: quiescence and outburst. However, at closer inspection, both stages cannot be completely separated. The presumed accretion disc and companion star, determining the nova behavior at quiescence, may also play a more or less important role during outburst due to their interaction with the ejected shell or the "bloated dwarf", which is formed in the explosion. On the other hand, the outburst stage gently merges into the quiescent stage, and it is possible that special features observed at quiescence are still influenced by effects from an outburst that lies more than 50 years in the past. It should be attempted to explain features in both stages with a similar model.

Due to the faintness of exnovae and the long time scales involved, the behavior of novae in the quiescent stage is only poorly known. Time scales for which observations are available are short compared with the whole interval of quiescence, which may last for ten thousands of years. Thus, our observations of classical novae are confined to a configuration shortly before outburst (and data of this stage are scarce, for obvious reasons) to a situation shortly (relatively speaking) after outburst. Therefore, statistical arguments and observations of some nova-like systems believed to be novae in quiescence in historical times must supplement our study of novae in quiescence. As will be seen in Section V.C, statistical arguments and the recent "hibernation" theory give arguments concerning

the possible identification of nova-likes or dwarf novae with novae in the middle of their quiescent stage.

This section will mainly deal with the discussion of models explaining the outburst behavior.

The concept of a nova "explosion" signals that we deal with a short-lived event, which seems to be determined by the phenomena that, according to our present understanding, occur in a thermonuclear runaway that lasts only minutes. In its aftermath, the outbursts of different novae show features that may not be completely determined at the moment of explosion. A closer look into the phenomenon shows that the outburst activity lasts a long time, at least weeks, if not years and decades.

I.B. THE LIGHT CURVE

The *optical* light curve shows a rapid rise, followed by a slower decline at variable rates for different novae. For some fast novae, oscillations occur in later phases; for some slow novae, they are found around maximum and later; a deep minimum is observed in the declining phase of some slow novae. In the *UV*, the maximum is delayed for smaller wavelengths. In the *IR*, excesses are sometimes seen after months, in general, coinciding with deep optical minima. In the *radio*, maximum occurs months after outburst and is delayed for large wavelengths. The *bolometric* light curve shows a much more gentle decline than the optical one. For some time, the nova radiates near the Eddington limit (or for fast novae, above it). A $M_{\max}-t_3$ -relation exists in the optical, where t_3 is the time in days to decline three magnitudes from maximum.

We first need to explain the light curve of novae, or, to be more precise, the variation of flux, integrated over all wavelengths, as well as its spectral distribution, with time. The variety of light curves observed for different novae (or even the differences between the light curves of two outbursts of a recurrent nova!) must be explained. The rise of the rate of nova energy production, the global optical decline, and the shift of the maximum of the distribution of radiation are compatible with a central object undergoing a thermonuclear runaway and suffering a strong mass loss, where matter is exhausted after a shorter or longer time. The detailed physics and the fine structure of a nova light curve are, however, hardly understood. In a few cases, brightness fluctuations with approximately the period of the underlying binary in early stages of the outburst are observed (V1500 Cyg, probably also V1668 Cyg). Repeated brightness (and possibly mass ejection) bursts, leading to secondary maxima and to short fadings and shifts of the energy distribution towards shorter wavelengths in the course of a prolonged maximum, are found in some novae (NQ Vul, QU Vul). Light oscillations in the transition phase occur mainly in a fraction of fast novae (GK Per, V603 Aql). Final declines show different gradients in the visual.

I.C. THE SPECTRAL EVOLUTION

The spectrum of a nova in outburst consists of superposed P Cyg profiles of different expansion velocities. In the early stages, the *blueshifted* absorption lines dominate; in intermediate stages, the spectrum has a typical P Cyg character; in the late stages, the continuum fades much more than the emission lines. The general evolution of the spectrum of a nova can be explained by the ejection of a large amount of mass, which starts at a well-defined time and declines in strength more or less slowly. This is indicated by the blueshifted absorptions, which are explained by the Doppler effect. The occurrence of lines of higher excitation and ionization in later stages of the outburst can be understood by the shrinkage of the photosphere, which becomes hotter at smaller radii if no large decrease in bolometric luminosity occurs. The gradual

thinning of the shell leads to the emergence of nebular lines, as will be seen in Sections II.D and IV.E. Complications occur, as can be noted from the observation of superposed P Cyg profiles of different expansion velocities, because of the interaction of material ejected at different times with different velocities (higher excitation lines, higher velocity components appear later.) The main bulk of matter, however, is ejected at the time of maximum: the line widths in the nebular stage correspond to velocity shifts observed in the principal spectrum. A velocity- t_3 -relation, and therefore a velocity-luminosity relation, exists.

An interpretation of the spectroscopic generalities derived from a comparative study of seven novae was carried out by McLaughlin (1943), where the spectral stages of the nova outburst are clearly described in much detail. A summary is found in McLaughlin (1942). This sequence was also adopted as a classification scheme by the IAU (Oort 1950). The systems listed in Table 7-1 seem to be normal in ordinary novae. Q indicates a spectral classification scheme, where each nova has a spectral Class Q, while subclasses are reached at various stages of the outburst (some may be skipped). A detailed description is found in Chapter 6.

I.D. THE EXPANDING NEBULA

For a small group of generally bright novae, expanding nebulae were observed in late stages. Structures at low resolution are, in general, spherical (contrary to jet-shaped). At high resolution or at later phases, they exhibit a wealth of detail: deviations from spherical symmetry are frequent; the shells appear often elliptical. The majority of shells shows the existence of equatorial rings and polar condensations. Additional structures are observed, e.g., the fragmentation of the shell into many cloudlets, or extended haloes.

Observational evidence shows that spherical symmetry for mass loss and velocity, often assumed in outburst models, is, in most cases, not correct. The observed ellipticity of the shell and the ring/blob structure may be connected with the underlying binary system, the fragmentation,

TABLE 7-1. SPECTRAL STAGES OF NOVAE

Spectrum	Absorption	Emission	Q
Prenova			0
Premaximum	I	I	1
Principal	II	II	2
Diffuse enhanced	III	III	3
Orion	IV	IV	4
N III (4640)	V	V	5
Eta Car = [Fe II]		VI	6
Nebular		VII	7
Wolf-Rayet		VIII	8
Final			9=0

with some sort of instability occurring in the ejected shell, the outer shell, with the survival of high-velocity material ejected in late stages of the outburst or with shock waves extending into the interstellar material.

The deceleration of nova shells observed in very late stages, decades after outburst, seems to be established. It is presumably caused by the interaction of the shell with the surrounding interstellar material.

It is often found that different regions of a nova shell have a different spectral appearance. A careful analysis must establish whether this is due to a different chemical composition, a different physical property (temperature, density), or, closely connected with it, different excitation conditions by the radiation field of the central object. This field probably has an axis of symmetry, but certainly deviates from spherical symmetry. Such "directional excitation" was suggested as early as 1937 by Grotrian in the case of DQ Her.

I.E. THE EXNOVA

Results of the few novae that have been observed at minimum are best interpreted in the framework of a close binary composed of a white dwarf and a low mass main sequence star. The nova explosion is a thermonuclear runaway in the electron degenerate hydrogen rich material close to the surface of the white dwarf, which was accreted (via a disk) from the late-type companion that suffers a Roche lobe overflow. It is usually assumed that the disk luminosity dominates all other contributions (white dwarf, late-type companion, hot spot) in classical nova systems at quiescence. Indeed, disk luminosities appear to have only little scatter if the different observational aspects are taken into account.

Recent detailed studies of the minimum visual magnitude of classical novae by Warner (1986, 1987) indicate that they have the same absolute magnitude at minimum, but are seen at different accretion disk inclinations. Warner showed that the histogram of minimum magnitudes can be explained by a mean absolute magnitude $[M_v(i$

$= 0^\circ) = 3.3]$, a random distribution of inclinations, and a scatter (intrinsic or observational) of 1^m . He argues that the minimum magnitude does not depend on nova speed class, or orbital period, or any other parameter besides the inclination i .

A study by Duerbeck (1986) indicates that minimum magnitudes do depend on speed class or light curve type, e.g., that slow novae (light curve Type D) are generally several magnitudes brighter than the rest. The very old novae CK Vul and WY Sge (Kenyon and Berriman, 1988) appear to be much fainter than an average, more recent exnova.

Spectroscopic studies of novae at minimum are only carried out for a very small percentage of the now identified objects at minimum. For some of them, radial velocity studies have been carried out. Preliminary results of a spectroscopic survey were reported by Duerbeck and Seitter (1987). The continua are frequently heavily reddened by interstellar extinction; weak to strong emission lines, mostly of the Balmer lines and notably He II 4686, are observed.

I.F. THE MANIFOLD OF NOVAE-SIMILARITIES, DIFFERENCES, AND THEIR POSSIBLE CAUSES

Despite the variety of light curve forms, spectroscopic development, and other features, there are similarities between different objects. Relations exist between the absolute magnitude at maximum and the t_3 -time and, less clearly, between the absolute magnitude at minimum and the light curve form, or between the velocity of expansion of the shell and the t_3 -time. This indicates that underlying systematics exist in novae. Possible "hidden parameters" that determine absolute magnitudes at maximum, ejection velocities, forms of light curves or shells are e.g.:

- The mass of white dwarf.
- The chemical composition of the white dwarf.
- The mass of accreted material.
- The accretion rate.
- The degree of mixing of white dwarf material into accreted hydrogen-rich shell material (or a

deviation of the composition of accreted material from the solar value).

- Orbital elements.
- The presence and strength of a magnetic field in the white dwarf.
- The inclination of the orbit.

Unfortunately, the statistics of data derived from minimum observations (orbital elements, properties of the binary components, and the shell) are poor, contrary to data on light curves and spectroscopic data obtained during outburst, or to minimum data for dwarf novae. Thus, hardly anything can be said about the importance of the above parameters for the characteristics or evolution of the outburst.

II. SHORT HISTORY

II.A. EARLY IDEAS TO 1930

Because of their sudden, spectacular appearance in the sky and their fading into oblivion, novae have been the object of speculation and theoretical approaches since earliest times. Quite often, the most advanced physical theories were used to explain the nova phenomenon. As an early example, Newton (1713) might be quoted, who, investigating the stability of cometary orbits, proposed that novae are old, burnt-out stars, whose supply of combustible material is replenished by the impact of comets. Another proposal that an aging star undergoes a nova event is due to Zöllner (1865). He thought that cracks in the surface of a cooling star allow magma from the interior to reach the surface, leading to a brightening. We see that novae were no more considered as "new stars" by some astronomers of the 18th and 19th centuries.

Theories of the nova phenomenon can be divided into two categories: one is the attempt to explain the underlying cause of the nova outburst; the other, the attempt to explain the photometric and spectroscopic features observed during outburst. We will first deal with the second question.

Increasing observational evidence, especially the availability of more and more detailed spectroscopic material, led to a reduction of the large number of explanations of the nova phe-

nomenon that were at hand in earlier centuries. When observing visually the spectrum of the (recurrent) nova T CrB, Huggins and Miller (1866) noted the appearance of bright and dark lines in the spectrum, without being able to note that the dark lines are blueshifted. They interpreted the bright H lines as a result of a gas explosion, a chemical combustion of hydrogen.

Blueshifted absorption lines were first observed in T Aur by Huggins (1892), Campbell (1892) and Vogel (1893), and shortly afterwards in IL Nor (Pickering 1894). Among the numerous collision scenarios, which were proposed to explain the two sets of lines of different radial velocity in nova spectra, the most influential one is that of Seeliger (1892, 1893, 1909). He assumed the penetration of a star into an interstellar gas or dust cloud, a scenario that won much favor after the light echo phenomena observed in the vicinity of nova GK Per (1901). The simultaneous occurrence of blueshifted absorption and stationary emission line as being caused by an extended, expanding envelope of gas was first proposed in a somewhat obscure way by Pickering (1894), then by Halm (1901) and finally by Lau (1906). However, in the early decades of nova spectroscopy, the blueshift of the dark lines was not unequivocally thought to be caused by the Doppler effect. A good survey of early theories is given by Stratton (1928).

The suddenness of the brightening, in combination with the different radial velocities of the dark- and bright-line spectrum, was always used as an argument for the collision hypothesis, i.e., the collision of a star and an interstellar cloud, two stars, a star and a comet, an asteroid, or a whole solar system. The extended pre-maximum stage of RR Pic and the postmaximum appearance of bright lines led Hartmann (1925) to the view that a single star "expands, explodes." In his second note, Hartmann (1926) states: "we conclude that the nova phenomenon is one which has its cause in the interior of certain stars. It is a disturbance of the physical-chemical equilibrium which occurs, without exterior cause, in a critical point of evolution, and leads to a rapid, explosive transformation of the whole celestial object. This

disturbance is possibly a radioactive transformation of atoms." Previous theories, especially the collision scenario, could not explain the delayed maximum of RR Pic and quickly fell into oblivion.

II.B. THE SEARCH FOR THE UNDERLYING CAUSE

After the external cause of a nova explosion had been discarded, and stellar structure became tractable, several outburst mechanisms were proposed. A good historical summary is found in Schatzman (1965).

Milne (1931) thought that the energy for the outburst is that of a star collapsing into a white dwarf. The energy is given (order of magnitude) by $\Delta E = G M^2 (1/R_2 - 1/R_1)$, which is 10^3 to 10^4 times higher than that observed. From the frequency of nova explosions, one can derive that a star must undergo about 10 nova explosions; i.e., a subset must show even more frequent eruptions. The hypothesis by Milne was relaxed by Vorontsov-Veljaminov (1940), who assumed that a sequence of nova outbursts of different strength leads to a transformation of a main sequence into a white dwarf.

Unsöld (1930), in investigating layers near the stellar surface for stability against convection, argued that the nova phenomenon is caused by the sudden onset of convection. One of his arguments was that the energy released in a nova explosion is much too small to be caused by dramatic changes in the interior, but he argued for a surface phenomenon.

Biermann (1939) thought that the explosion occurs because of an instability in a hydrogen-poor subdwarf. An instability develops at a certain depth where radiant equilibrium changes over to adiabatic equilibrium. In this case, the star should be poor in H and He, and the changes occurring lead to a release of ionization energy.

In 1946, Rosseland (1946), Lebedinskii (1946) and Schatzman (1946) thought of the nova explosion as a shock wave in the stellar interior, and Rosseland came up with some sort of atomic bomb explosion, specified later by Schatzman

(1950, 1951), who postulated as the energy source the conversion of ^3He into ^4He . Gurevich and Lebedinskii (1947a,b) assumed that a nova explosion results from explosive low-temperature nuclear reactions taking place in the peripheral layers of a star, which produces energy by the CNO cycle.

Mestel (1952) examined accretion of interstellar matter onto hot white dwarfs and postulated either continuous burning or supernova-type explosions.

The discovery by Walker (1954) that the famous nova DQ Her is a close binary was soon generalized by Struve (1955) and Huang (1956) to the possibility that all novae and nova-like stars are binaries. Kraft (1963, 1964) came up with the still-favored model: that the nova is a close binary composed of a white dwarf and a low-mass main sequence star. The nova explosion is a thermonuclear runaway in the electron degenerate hydrogen-rich material close to the surface of the white dwarf, which was accreted (via the accretion disk) from the late-type companion overflowing its Roche lobe. The binarity and the inclusion of an accretion disk led nova research out of a dilemma that still intrigued Payne-Gaposchkin (1957): novae at minimum are blue objects at $M_V \approx +4$; this always leads to objects below the main sequence, but above the white dwarf region, i.e., to subdwarfs, similar to the central stars of planetary nebulae.

An early nova model including a close binary is due to Schatzman (1958). He suggested that in one of the components, nonradial forced oscillations are produced by orbital motion. If there is a resonance between the orbital period and one of the periods of nonradial oscillation, the amplitude of the force oscillation is finite unless the damping constant vanishes or is negative. When this happens (as a consequence of the secular evolution of the star), the forced oscillation may become an explosive process. The damping constant can vanish if the energy sources are close enough to the surface of the star. One can predict the ejection of matter along polar caps and one or several belts or zones. The pole of the system is in the direction of the perturbing star. This is one of

the first models (and nearly the only one) that make predictions about the directional dependence of mass loss.

In the 1960s, a clearer understanding of the structure of the binary components in nova systems led to a more realistic scenario for nova explosion: A Roche-lobe overflowing, unevolved star is losing H-rich material, and a fraction, if not almost all, of this material is accreted by the white dwarf companion. As accretion continues, a layer of H-rich material is built up on the surface of the white dwarf, and the bottom of this layer will be gradually compressed and heated until it reaches ignition temperatures of H-burning reactions. Hydrodynamic studies [Giannone and Weigert (1967), Rose (1968), Starrfield (1971a, 1971b)] have shown that a thermonuclear runaway will occur in the H-rich envelope, which can produce the observed energies and can reach temperatures and energy-generation rates at which dynamic effects become important (Starrfield 1971a, 1971b).

II.C. MECHANISMS OF MASS EJECTION

Despite the fact that the energy source for the nova explosion remained obscure for a long time, different mechanisms for mass ejection were investigated: shock ejection, pressure ejection, radiation pressure, and pulsational mass loss. It is now believed that the first three can act during various stages of the nova outburst. While shock ejection seems to play a predominant role in fast and recurrent novae, thick stellar wind driven by radiation pressure is important for slow novae.

II.C.1. MASS EJECTION BY EXPLOSIVE EVENTS

Rosseland (1946) proposed a central explosion (due to some "subnuclear mechanism"). The observed phenomena (several maxima) can be interpreted by the effects of a shock wave reaching the surface. Shock ejection was also proposed by Lebedinsky (1946). Schatzman (1946, 1949) also assumed that mass loss of novae is due to the propagation of a shock wave.

When energy is generated in a region in a time interval that is shorter than the sound travel time across that region, a shock wave is generated at the bottom of the region, which propagates outward. As the wave moves outward into less dense material, the shock wave accelerates. After the passage of the shock, the ejected material is left with a steep velocity gradient (Figure 7-1), see also Hazlehurst, 1962; Sparks, 1969). Shock ejection mechanisms lead to a temporal behavior of mass loss velocities, which shows a marked decrease of speed. In general, this is contrary to observations of classical novae, except perhaps for material of the premaximum absorption system.

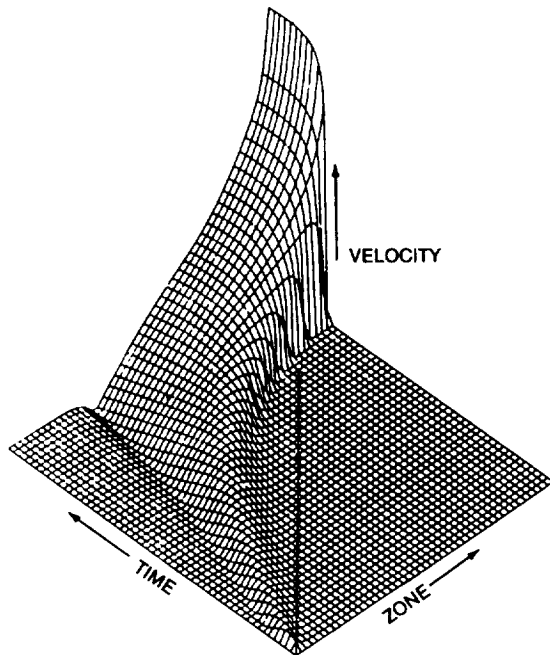


Figure 7-1. A three-dimensional plot of the velocity of material as a function of time in various zones of a star for shock ejection. The increase of the shock velocity and strength and the steepening of the shock front can be seen, as the shock wave propagates outward to zones of lower density. After the shock wave has passed through, the material has a large differential velocity. The trough on the left is due to material which did not reach escape velocity and is falling back onto the star (Sparks, 1969).

II.C.2. PRESSURE EJECTION

If the energy is deposited in a region during a time interval that is longer than the sound travel

time across that region, but shorter than the time interval of radiation diffusion of the region, a "pressure" wave develops, which ejects all the material above the region with roughly the same velocity; i.e., the velocity gradient is much smaller than in the case of shock ejection (Figure 7-2). This mechanism was investigated in detail by Sparks (1969).

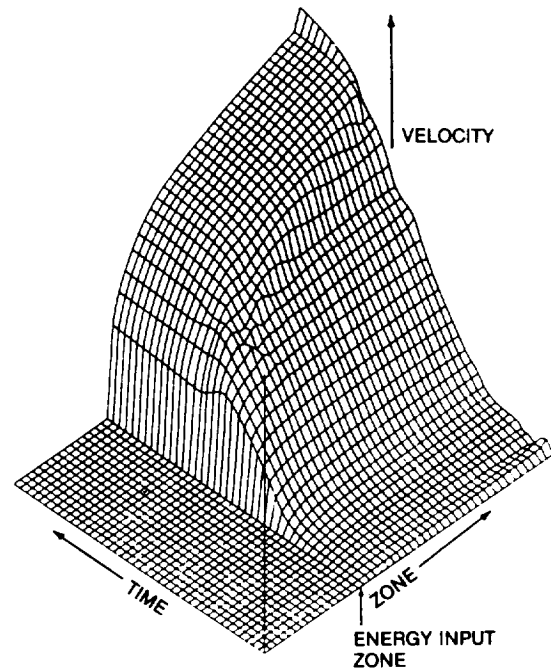


Figure 7-2. A three-dimensional plot of the velocity of material as a function of time in various zones of a star for pressure ejection. A pressure front is formed which ejects all material above the energy-input zone with roughly the same velocity. The slight increase in the velocity of the outer zones is due to the outward-propagating shock wave (Sparks, 1969).

II.C.3. RADIATION PRESSURE

The importance of radiation pressure by electron scattering for stars of high luminosity was pointed out by Eddington (1921). Radiation pressure in spectral lines was studied by Milne (1926), who derived limiting velocities of the order of 1600 km s^{-1} . It was further discussed by McCrea (1937), and in more recent times by Friedjung (1966a,b,c), Finzi (1973), and Nariai (1974), and in the optically thick wind models by Bath and Shaviv (1976) and Bath (1978).

II.C.4. PULSATION MASS LOSS

Problems encountered with shock ejection led Rose (1968) to investigate pulsational instabilities. Assuming thermally unstable H burning, a hot white dwarf becomes pulsationally unstable. The dissipation of pulsationally produced shock waves appeared to be a plausible means of surface heating; however, the calculations do not show how this extended envelope is ejected (Rose and Smith, 1972). In a later study, Sastri and Simon (1973) investigated multimodal radial pulsational instability in a prenova model. Predictions to test these models are, however, lacking.

II.D. CONTINUOUS EJECTION

First ideas on continuous ejection and the existence of a photosphere that shrinks in the later stages of the outburst, while the temperature increases, are found in Halm (1904) and Pike (1928, 1929). Whipple and Payne-Gaposchkin (1936) applied the theory of continuous ejection to DQ Her. These early studies found it difficult to distinguish observationally the process of continuous ejection from the expansion of a shell, both processes suggested already by Halm (1904). The process of continued ejection is favored by Whipple and Payne-Gaposchkin for the following reasons: if the continuous spectrum and the absorption lines are both produced in an expanding shell, the smallness of the observed changes in the absorption spectrum and in the energy distribution of the nova in its rise to maximum would require a continuous balance between the total radiation and the radius of the initial shell, unless one considers the shell so thick that it dams back the radiation for a considerable time. If the shell were as thick as this, one should expect the absorption lines to be broadened, because they would be formed near the surface of the rising photosphere. On the other hand, the observed narrowness of the lines indicates that the absorption must be produced so far above the effective photosphere that only a small solid angle is subtended at the center of the star by the material producing the absorption spectrum. For a thin shell, the intensity of the absorption lines might be expected to decrease with time but observations of DQ

Her show that the intensity of the H lines increased.

II.E. SUPER-EDDINGTON LUMINOSITY

Finzi (1973) showed that the radiant energy of a nova in the course of its outburst cannot be stored and released by the ejected envelope, and concludes that the energy radiated in a nova outburst, i.e., 10^{45} erg, is released after the explosion by the central star. The luminosities of some novae at maximum exceed the "critical" or Eddington luminosity

$$L_{\text{CRIT}} = \frac{4\pi c G M_{\odot}}{\kappa_{\text{TH}}} \frac{M}{M_{\odot}} = 6.5 \times 10^4 \frac{1}{1+X} \frac{M}{M_{\odot}} L_{\odot}$$

where X is the relative H abundance in the external shell of the nova and $\kappa_{\text{th}} = 0.2 (1 + X)$ is the Thomson opacity, the opacity of fully ionized matter at low density, $c = 3 \times 10^{10}$ cm/s, $G = 6.67 \cdot 10^{-8}$ dyn cm²/g², and $M_{\odot} = 2 \cdot 10^{33}$ g. The corresponding limit to the absolute bolometric magnitude is about $-6.8 - 2.5 \log_{10} (M/M_{\odot})$ for $X \approx 0.7$. The critical luminosity is an upper limit to the luminosity of a star in hydrostatic equilibrium (Eddington, 1921). Finzi found that the luminosities of all postnovae (at least shortly after outburst) are larger than the critical luminosity and that their photospheres are steadily flowing out. Below this radiative atmosphere, Finzi assumed a convective hydrogen-rich shell. The concept of novae radiating above or near the Eddington luminosity plays an important role in subsequent nova wind models by Bath and Shaviv (1976) and Bath (1978).

II.F. HISTORY AND RESULTS OF MODELING (FROM 1930)

Grottrian (1930), and Menzel and Payne (1933) found evidence, mainly from the appearance of forbidden lines, that diminishing pressure, rapidly rising temperature of the region in which the ionizing radiation originates, and dilution of radiation can account for the order of appearance of spectral lines of successive stages of excitation and ionization, as well as the appearance of lines originating by transitions from metastable levels.

The interpretation of observational results of nova DQ Her (1934) led to a new approach to the question of the time interval during which shell ejection occurs. Practically at the same time, and likely independently, Gordeladse (1937), Grotian (1937) and Whipple and Payne-Gaposchkin (1936) put forward models with continuous mass outflow to explain some of the features of the slow nova DQ Her.

McLaughlin (1943) pictured the rise to maximum as an eruption of a spherical shell of gas sufficiently dense and deep to be opaque and to behave like an expanding star. The photosphere is only an optical level in the outward-rushing cloud, and as it expands, the individual atoms migrate from subphotosphere through the photosphere to the reversing layer. Finally, at maximum light, the cloud, which is now detached from the star, suddenly becomes transparent. In the early postmaximum stage, there are conspicuous bright bands that originate in gases located all around the star. The emission originates, however, mainly in the inner layers of the shell—almost on its inner surface—while absorption is produced throughout a great depth.

McLaughlin thought that the rate of mass ejection shows a variation similar to and almost in phase with the light curve, anticipating the latter very slightly. The model involves a main burst, followed by continuous expulsion of matter at a steadily decreasing rate. It requires a sharp reduction of the rate at the end of the main burst. The mass loss then drops slowly and continuously.

Qualitatively, this model has expansion accompanied by cooling. The shell acts initially as an (expanding) photosphere and mimics, after a short time from outburst, a spectral type of late B or early A with extended atmosphere. In a short time, τ_{vis} in the shell becomes less than unity; the shell becomes optically thin and ceases to radiate like a photosphere. The visual continuum fades and its color temperature increases. Further expansion of the shell produces

a. Less absorption of stellar radiation, thus transformation of P Cyg profiles into emission

profiles.

b. Dilution of radiation, with accumulation of atoms in metastable states.

c. Time intervals between collisions greater than the lifetimes of metastable states: forbidden lines can be formed.

There is also formation of a subsequent system of shells, closer to the star, that are explained as being due to the tail of the ejection rates after the main burst (see Figure 7-3a and 7-3b). McLaughlin's model is indeed one with continuous ejection over a period of years (1943, p.188 ff.)

On the other hand, Pottasch's model (1959 a,b,c,d) is based on a spherically symmetric main burst only. Thus it offered the possibility to be tractable. The following assumptions were made:

a. The geometrical thickness of the shell is small compared to the radius of the shell.

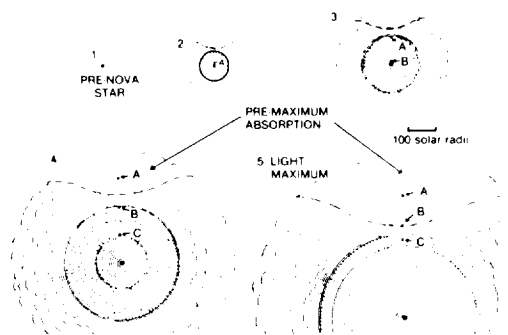


Figure 7-3a. Cross sections of a nova during the rise to maximum, as envisaged by McLaughlin (1950). The observer is looking from the top of the figure. The large black dot is the main body of the star. Stippled area represents the densest part of the ejected shell, concentric circles represent the optically thinner photospheric layers, which merge into the cloudy forms that represent the true atmosphere. On each drawing, a heavy dashed line outlines the region that is effective in producing the observed absorption spectrum. Successively ejected atoms A, B, and C are shown. With expansion, the layer containing A, and later that containing B, become transparent. By light maximum, the ejection has diminished and the shell, still opaque, has become detached from the star.

b. The density of the shell is always uniform and varies inversely as the volume of the shell.

c. The photosphere of the star remains at the

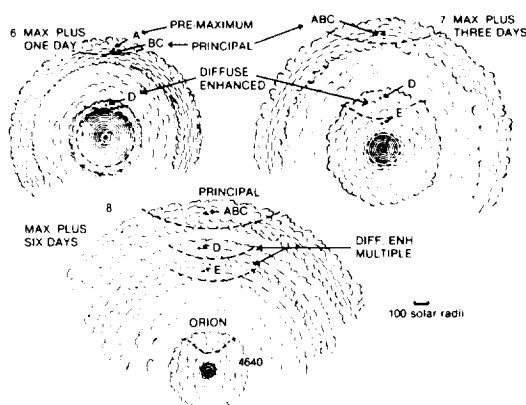


Figure 7-3b. Cross sections of a nova during early decline, as envisaged by McLaughlin (1950). Symbols have the same meaning as in Figure 7-3a. In section 6, the shell has just become transparent. Atoms B and C, owing to acceleration of the inner layers, have overtaken A, and the resultant shell is the principal one, the premaximum spectrum having disappeared with the engulfment of the outer layers (A) in the accelerated inner ones (B and C). Atom D is contributing to the diffuse enhanced absorption. In section 7, the inner cloud has become more extensive and the diffuse enhanced absorption is correspondingly stronger. Atom E is entering the region of absorption. In section 8, the gas ejected in the diffuse enhanced stage is in the form of two detached shells, overtaking the principal shell, while the inner cloud has developed into the Orion and the 4640 spectral stage.

initial temperature of the shell. As a result the inner surface of the shell receives stellar radiation diminishing as the inverse square of the time.

d. Energy is transferred by radiative processes.

A numerical solution was carried out for the following conditions: shell mass $M = 5 \cdot 10^{-3} M_{\odot}$, velocity = 800 km/s, central star temperature = 200,000 K, and radius $0.3 R_{\odot}$. The computation was carried out until the shell has an optical depth low enough to question the basic assumptions. The model reproduces the essential features of a (fast) nova light curve with even the premaximum halt.

Pottasch's results are

1) The absolute magnitude of about -8.0, which the light curve attains before the shell becomes optically thin, is reasonable for a fast nova.

2) The temperatures observed during days 2 and 3 are similar to those observed.

3) The initial "bump" may be identified with the premaximum halt, the physical reason being that the shell is assumed to be initially isothermal at the surface temperature of the star from which it came, and the shell initially cools faster than the radiation can be transferred through the shell. Thus, the initial loss of energy occurs quickly and then subsides. The transfer of the energy of the star to the surface occurs more slowly; the combination of the two curves has a bump (more details are given in section III.A).

II.G. VELOCITY GRADIENTS

A plot of measured radial velocities of absorption versus time tends to increase, also when the feature is attributed to the same "system," which may form in a given layer. Furthermore, later spectral systems (diffuse-enhanced and Orion) always show larger radial velocities of absorption lines or wider emission lines. Thus, different parts of the shell move with different velocities. The diffuse-enhanced and Orion systems are attributed to the inner region surrounding the nova, while the original emission and absorption systems belong to the principal shell, which has moved outward in the meantime. Thus, we expect interactions with previously ejected material.

Arguments for the "inside origin" of the diffuse-enhanced and Orion feature have been outlined by McLaughlin (1947):

1) Some novae have secondary light variations, and the diffuse-enhanced absorption undergoes marked changes of intensity and of structure and position related to the light variations, while the principal spectrum responds only slightly with changes of intensity or excitation, but without changes of displacement.

2) The broad emissions of the diffuse-enhanced system extend across the emission and

absorption of the same lines from the principal shell. Nevertheless, the principal absorption remains strong and well-defined, without the filling-in that would occur if the atoms that produce the diffuse-enhanced spectrum were outermost.

3) There are numerous examples of partial or complete obliteration of absorption lines of the diffuse-enhanced spectrum by overheating emissions of the principal spectrum (further discussed in Section IV.A).

4) The same arguments apply also to the Orion spectrum.

According to McLaughlin, the high-speed atoms must eventually overtake the principal shell. Those that produced the Orion spectrum form a haze so rarefied that it would probably have no observable effects. Those of the diffuse-enhanced spectrum are present early enough to overtake the principal shell, while absorptions are still distinct. The collision of clouds should cause sudden disappearance of components of the diffuse-enhanced system, the appearance of new components, and acceleration of the principal shell. The principal shell is probably more massive than all the matter ejected later, so that large accelerations by collision are not to be expected. The principal shell continues to move outward until it becomes a quasi-planetary nebula. After a few decades, this has become too faint to be observable.

II.H. DETERMINATION OF T_{col} , T_{ion} , T_{exc} , T_e , N_e

The appearance of the bright nova DQ Her, in 1934, marked a decisive point in the study of novae, since at that time, the tools of stellar and nebular diagnostics were already well developed. Color temperatures T_{col} were most extensively determined (see McLaughlin, 1960, for a summary). Ionization temperatures T_{ion} were determined using Zanstra's (1931) method; the first to use it was Beals (1932) for V603 Aql, later determinations were made for CP Lac (McKellar, 1937) and DK Lac (Larsson-Leander, 1953,

1954). As a general rule, ionization temperatures are systematically much higher than color temperatures, and values based on He II and N III are approximately double those of H, while the nebular lines give values somewhat lower than hydrogen.

Excitation temperatures T_{exc} are calculated mainly from the relative intensities of [O III] and H lines (Stoy, 1933) and from the ratio He II 4686/H β (Ambarzumian, 1932). Both methods yield temperatures near or larger than the Zanstra temperatures. Oehler (1936) applied these methods to the nebular spectrum of DQ Her.

Electron temperatures T_e traditionally were mainly derived from the ratio of the [O III] lines (5007 + 4959)/4363. For all novae in which the ratio has been measured, T_e ranges between 6000 and 10,000 K, with a tendency to decrease as the nova fades (gain of strength of the nebular lines over the auroral transition 4363). Early applications to novae were made by Popper (1940) and Gaposchkin and Payne-Gaposchkin (1942).

The electron density is derived from the linear size of the nebula, and the surface brightness (or flux) of a hydrogen line with negligible self-absorption (Ambarzumian, Kosirev, 1933; Sayer, 1940; Whipple and Payne-Gaposchkin, 1936; Payne-Gaposchkin and Gaposchkin, 1942), which turned out to range from 10^9 cm^{-3} in the early postmaximum stage to 10^6 cm^{-3} in the early nebular stage. With the assumption of $N_e = N_{\text{H}}$, the density of hydrogen ions, total nova masses were calculated by the above-mentioned authors.

II.I. ABUNDANCE DETERMINATIONS BY CURVE OF GROWTH OR NEBULAR LINES

The similarity of some nova spectra at maximum with those of supergiants led Mustel and Boyarchuk (1959) to attempt coarse analysis of nova spectra to determine excitation temperatures, microturbulent velocities, and chemical abundances. A number of analyses, mainly on slow novae, were carried out in subsequent years.

Another way of determining abundance of nova shells is by analyzing their nebular emission line spectrum. A wide range of ionization conditions is commonly observed in nova envelopes. Because of the large ejection velocities, emission lines are often wide, making line identifications sometimes problematic. The procedure employed to determine the chemical composition of nova shells is the same as that employed in abundance analyses of planetary nebulae. The first large-scale analysis of nebular spectra of the novae V603 Aql, RR Pic, GK Per, CP Lac, and DQ Her was carried out by Pottasch (1959e). He found that the abundances of O, N, S, Ca, and Ne seem to be a factor of 5 greater than cosmic.

III. SIMPLE MODELS TO EXPLAIN OBSERVATIONS

Various simple models to explain the observed light and spectral observations during post optical maximum activity are conceivable. Such models were described by Friedjung (1977a). They all describe stars that eject high-velocity gas during a limited time and that have a temporary increase in brightness. The geometry and the kinematics differ from model to model. As will be seen, the true situation is more complex, and though one simple model may be more helpful in explaining many phenomena, others may be needed to interpret other aspects. However, each of these models is conceptually very useful. Five of these models will be described.

III. A. INSTANTANEOUS EJECTION I

In instantaneous ejection models, all or nearly all material is ejected in a time that is short, compared with the duration of postoptical maximum activity. The observed changes occur in previously ejected gas, which at first is optically thick in the continuum (an expanding atmosphere is seen), and which later becomes optically thin. Instantaneous ejection type I models are those where the ejected material is in a fairly thin shell, the thickness of which remains small. Supposing that the outer radius of the shell at optical maximum is r_0 and its thickness is Δr_0 , while the corresponding values at a time near the end of activity are r_1 and Δr_1 , instantaneous ejection type I sup-

poses that, as $r_0 \ll r_1$,

$$\Delta r_0 \text{ and } \Delta r_1 \text{ both are } \ll r_1 - r_0 \approx r_1 \quad (7.1)$$

To satisfy this condition, it is necessary for the expansion velocity of the shell to be much larger than the velocity of increase of thickness of the shell.

Assuming spherical symmetry and density of the shell to vary as $1/\text{volume}$, this density will vary as $r^2/\Delta r$, r and Δr being the shell radius and thickness at any time. For a constant expansion velocity, the density will vary as t^{-2} at time t from outburst, as long as Δr is constant; it will vary as t^{-3} if Δr increases with a constant velocity from a value of zero at time zero. The shell surface area will vary as t^2 for a constant expansion velocity.

It can be supposed that the ejected shell is optically thick in the beginning; its surface area is then that of an expanding photosphere. After optical maximum, the bolometric luminosity does not increase, so if instantaneous ejection Type I with an optically thick shell were true, one would expect the temperature of the latter to fall. Once it became optically thin, the central object would become visible. The emission measure at this later stage varies as $r^2/\Delta r$, so for a constant expansion velocity and shell thickness, it varies as t^{-2} ; when Δr also increases from zero with a constant velocity from time zero, the emission measure varies as t^{-3} .

Instantaneous ejection Type I was extensively studied by Pottasch (1959 a,b,c,d). The shell thickness was supposed to increase from zero at ejection with a velocity equal to $2a/(\gamma - 1)$, where a is the speed of sound and γ , the ratio of specific heats equal to $5/3$. This is a theoretical rate of expansion into free space.

In spite of the shortcomings of the model, which will be described later, Pottasch had some successes in explaining observations. First, he was able to explain the shape of the light curve before optical maximum including the premaximum halt, taking account of the fact that in early stages the optically very thick shell should not

be in radiative equilibrium (Pottasch 1959b). Moreover, this result appeared in a calculation for a shell mass of 10^{29} g ejected at 800 km s^{-1} surrounding a central star with a constant temperature of $2 \cdot 10^5 \text{ K}$ and radius of $2 \cdot 10^{10} \text{ cm}$ (Figure 7-4). In Pottasch (1959c), early theory (largely neglecting velocity fields) was used to calculate the flux of the $\text{H}\alpha$ line, when it was optically thick, as a function of electron density. He used the thickness of the shell given by this model to compute its mass, which could be compared with a mass calculated in late stages when the shell is expected to be optically thin

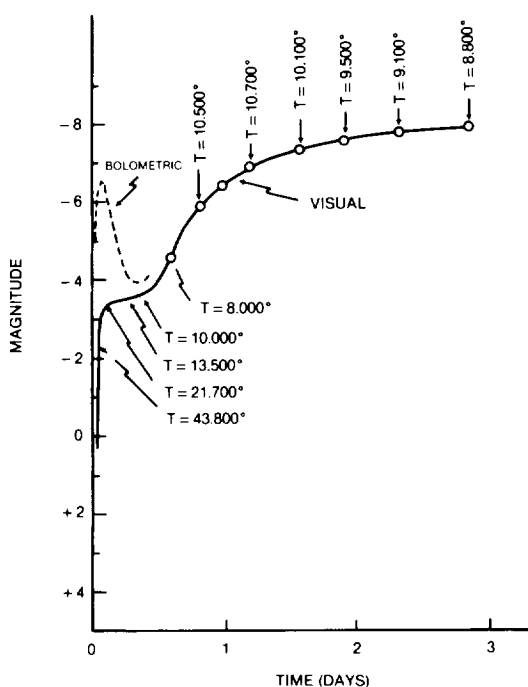


Figure 7-4. Light curve of the expanding nova shell according to computations by Pottasch (1959b). The energy emitted by the shell is shown as a continuous line as visual magnitude, the broken line gives the bolometric magnitude. The surface temperature is given at different stages. Note the presence of a pre-maximum halt in the visual light curve.

in the Lyman continuum (all hydrogen ionized) and in $\text{H}\alpha$. In spite of the various assumptions and approximations (including, for instance, that the $\text{H}\alpha$ flux was three times that of $\text{H}\beta$, which unlike $\text{H}\alpha$ was observed in all the studied novae), order of magnitude for the two masses determinations was reached for four out of five novae.

It may be noted that even Pottasch's (1959d) calculations indicate that part of the story was missing from his model. Temperatures of the central objects were found from the ratio of line emission emitted by ionized helium to that emitted by hydrogen, assuming photoionization by a Planckian continuum. The radius of the photosphere could then be determined from the total flux of photons emitted shortwards of 912 \AA , taken as equal to the total flux of Balmer emission photons. The calculated photospheric radii, for dates as soon as 10 days after ejection of the postulated shell, were much smaller than the latter, which for consistency needed already to be optically thin longwards of the Lyman limit. In addition, these radii decreased with time, suggesting that a constant radius "nova remnant" was not seen. It is this varying central object that first suggested that another sort of model needs to be invoked.

III.B. INSTANTANEOUS EJECTION TYPE II (SOMETIMES CALLED "HUBBLE FLOW")

In this model, a thick envelope is ejected instantaneously. This envelope remains thick as different parts have different velocities. In the simplest situation, where the velocity of any particular mass remains constant, the distance travelled by it at a particular time is proportional to the velocity. Thus, the outermost parts of the envelope have the largest velocity and the innermost parts, the smallest one. As expansion occurs, the optical thickness decreases, and the radius below which the deeper layers are not visible (which equals the photospheric radius) shrinks. It can be much smaller than the envelope.

When different parts of the envelope have the same density and each part a constant velocity, the density will vary as t^{-3} and the emission will measure as t^{-3} , t being the time since ejection. This behavior is the same as that of a thin shell whose thickness uniformly increases from zero at the time of ejection.

In the framework of instantaneous ejection Type II, one can imagine a much higher density in inner than in outer parts of the envelope. In this case, inner regions could remain optically thick for a

long time. However, they would produce strong line emission, because emission due to recombinations (and also collisional excitation) is proportional to the square of the density. As the radius of the photosphere decreases, regions with lower and lower expansion velocities would become visible, and the emission line profile would be more and more dominated by slowly expanding material. One might then expect the FWHM of emission lines to decrease with time, and the violet shift of P Cyg absorption components to decrease as well.

The type of situation just described is believed to be true for supernovae. Violet shifts of absorption lines have been seen to decrease in Type II supernova spectra (Chugai, 1975). Supernova 1987A in the LMC is a very good example (For instance, see Hanuschik and Dachs, 1987, and Henbest, 1987). The model has been reasonably successful. However, this situation does not seem to be true for nearly all classical novae, except perhaps before optical maximum, when the premaximum system can have behavior characteristic of instantaneous ejection Type II.

In spite of this, such models have sometimes been suggested for classical novae, including, for instance, Sobolev (1960). Nariai (1974) calculated the position of the photosphere, taking into account gravitational deceleration. In early stages, when the radius of the photosphere is large and the effective temperature is low, hydrogen is not ionized, and the radius of the photosphere is predicted to vary as $t^{0.7}$. At a later stage, hydrogen is predicted to become ionized and the photospheric radius, to decrease very rapidly, as shown in Figure 7-5. Instantaneous ejection Type II or a "Hubble flow" was invoked by Seaquist and Palimaka (1977) to explain the radio emission of FH Ser. It may be noted that the regions from which radio emission is detected are much larger than those of the optical and ultraviolet photosphere (because of the free-free absorption coefficient), thus a simple model as the one given here gives a better description for the radio observations, as compared to optical and ultraviolet observations.

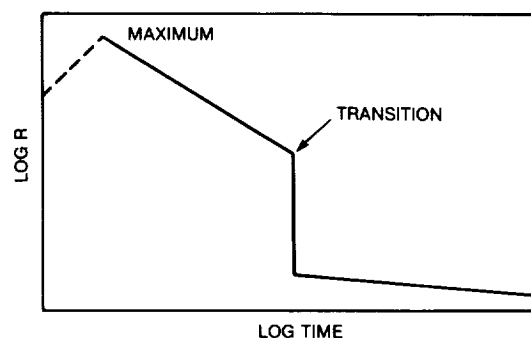


Figure 7-5. This schematic diagram gives the position of the photosphere, assuming constant luminosity, according to Nariai (1974). His study describes the slope of the curve from maximum to the transition phase, and also the sudden decline. After transition, the light curve is produced by the surface of the remnant in quasi-static contraction.

III. C. CONTINUED EJECTION A

Continued ejection models emphasize the importance of winds from the nova after optical maximum. When continued ejection A occurs, most of the emission of the continuous spectrum in the optical and ultraviolet comes from an optically thick wind. The photosphere (sometimes called quasi-photosphere) is located in the wind. The radius of this photosphere is directly related to the mass loss rate. The optical fading is readily associated with a drop in this mass loss rate, leading to a smaller photospheric radius, and if the bolometric luminosity does not decrease very rapidly, the photosphere will become hotter as it shrinks. The photosphere is much smaller than the ejected envelope for continuous ejection A, except near optical maximum. Material ejected near optical maximum will produce what is later an optically thin density peak near the outer edge, while the density will also be high in the inner regions containing material that has not expanded very much. This kind of model for instance, has been supported by Whipple and Payne-Gaposchkin (1936), Friedjung (1966 a,b,c), Bath and Shaviv (1976), and by Bath (1978).

Early quantitative formulations of continued ejection used the theory of Kosirev (1934) and Chandrasekhar (1934) for extended grey atmospheres in Local Thermodynamic Equi-

librium (LTE). Simple opacity laws were assumed such as a constant opacity (electron scattering dominant) or a mean photoelectric opacity law. As suggested by observation, the time scale of variations was supposed to be long, compared with the time for ejected material to travel from the mass losing star to the photosphere; therefore, the density in optically thick regions was supposed to vary as r^{-2} , r being the radial distance from the centre.

Making an Eddington approximation, one obtains

$$B = \frac{3}{4} \pi F r^2 \int_0^{\tau} \frac{d\tau}{r^2}, \quad (7.2)$$

where B is the intensity of black body radiation emitted at a distance r from the centre, πF is the radiation flux, and τ is the optical depth. In LTE,

$$\pi B = \frac{1}{4} a c T^4, \quad (7.3)$$

with T being the temperature. Supposing that the mean absorption coefficient is given by

$$\kappa = \kappa_0 P_e T^{-n}, \quad (7.4)$$

with P_e the electron pressure, and that the number of free electrons per nucleus is a constant $1/\alpha$, one obtains that

$$T \text{ varies as } r^{-n/(n+3)} \quad (7.5)$$

In an electron scattering case, κ is replaced by a constant σ , and

$$T \text{ varies as } r^{3/4} \quad (7.6)$$

Having obtained a temperature law, one can calculate the flux that should be observed at each wavelength by integrating emission from different directions at different optical depths, if one supposes that local emission is always Planckian. It may be noted that in regions where electron scattering dominates, this would not be correct even if all other hypotheses were valid. The result of such calculations

is a way to interpret observations of continuum flux and color temperature in terms of photospheric radii and effective temperatures. Put in another way, one can correct temperatures and radii determined, assuming that the photosphere has a Plankian energy distribution.

If one knows the photospheric radius and temperature, the mass flux can be calculated. Putting the outflow velocity equal to V (assumed constant), the photospheric radius and temperature equal to R_p and T_p , respectively, and assuming for a first approximation an optical depth at the photosphere of $2/3$, the mass loss rate is

$$\dot{m} = \sqrt{\frac{32(2n-14)}{3(n+3)}} \frac{\mu(\alpha+1)}{R K_0} \pi V T_p^{(n-1)/2} R_p^{3/2}, \quad (7.7)$$

(photoelectric κ)

or

$$\dot{m} = \frac{8\pi}{3} V \frac{\mu\alpha}{\sigma} R_p, \quad (\text{electron scattering}) \quad (7.8)$$

Here R is the ideal gas constant, and μ the molecular weight.

This highly simplified theory illustrates some important features of continued ejection A. The characteristics of the continuous spectrum are directly related to the mass loss rate. If fading is rapid, R_p and \dot{m} decrease rapidly. The time variation of the ejection rate derived from a good theory of an optically thick wind can be compared with other constraints to test continued ejection A. Such constraints, as shall be seen later, are connected with collisions between different parts of a nova envelope moving at different velocities, accelerating, for instance, the slower material.

Bath and Shaviv (1976) considered continued ejection A from a slightly different point of view. They supposed that during post optical maximum activity, the luminosity is close to the Eddington limit. The luminosity needs to reach this limit for radiation pressure to accelerate an optically thick wind, while a larger luminosity would lead to a breakdown of hy-

drostatic stability of the central mass losing star. In the case of constant luminosity L ,

$$T_p = R_p^{1/2} (L/\pi\sigma c)^{1/4}, \quad (7.9)$$

with T_p the photospheric temperature, taken as equal to an "effective temperature" for the present approximation. The density in the photosphere ρ_p is given by

$$\rho_p = \sqrt[2]{\frac{2\mu(\alpha+1)(2n-14)}{3RK_O(n+3)}} T_p^{(n-1)/2} R_p^{-1/2},$$

(photoelectric κ) (7.10)

or

$$\rho_p = \frac{2}{3} \frac{1}{\sigma R P}. \quad (\text{electron scattering}) \quad (7.11)$$

Combining Equation (7.9) with Equation (7.10) or (7.11), one sees that if L is constant, there are one to one correlations between photospheric density, temperature, and radius. These physical conditions are for constant L , which if LTE is valid, is also correlated with visual luminosity. In this way Bath and Shaviv (1976) were able to explain the frequent correlations of physical quantities with visual brightness during the decline of a nova from visual maximum.

Other models including continued ejection A were developed. Bath (1978) improved opacities. When electron scattering dominates (as is generally the case), he assumed an effective opacity of $(\kappa\sigma)^{1/2}$ in our notation, κ being a Cox and Steward opacity, quoted by Bath, as taking account of the fit and extension due to Christy (1966). The difference between scat-

tering and absorption was thus taken into account, and some results are shown in Figure 7.6 Harkness (1983) calculated radiative transfer rigorously for non-grey optically thick winds; results are shown in Figure 7-7. However, these winds were not only in LTE, but they also had a solar composition and a constant luminosity at the Eddington limit. Such assumptions are at least partially wrong, as shall be seen.

Friedjung (1966c) showed that the optically thick winds required for continued ejection A could be accelerated by radiation pressure at very large optical depths. More detailed hydrodynamic calculations were performed by Ruggles and Bath (1979), while Friedjung (1981) studied observational consequences for acceleration of a wind for a luminosity far above the Eddington limit. We shall return to these questions later.

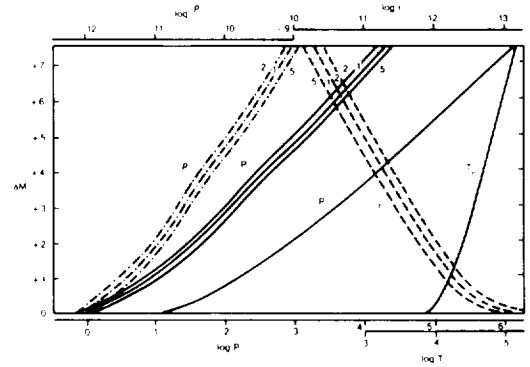


Figure 7-6. Variation of photospheric density, pressure, radius and temperature for outflowing nova winds as a function of ΔM_v , the decline in visual magnitudes, for a blackbody continuum at constant luminosity in a nova model of Bath (1978). Three luminosities, $0.5 L_{ed}$, L_{ed} , and $2 L_{ed}$, are considered.

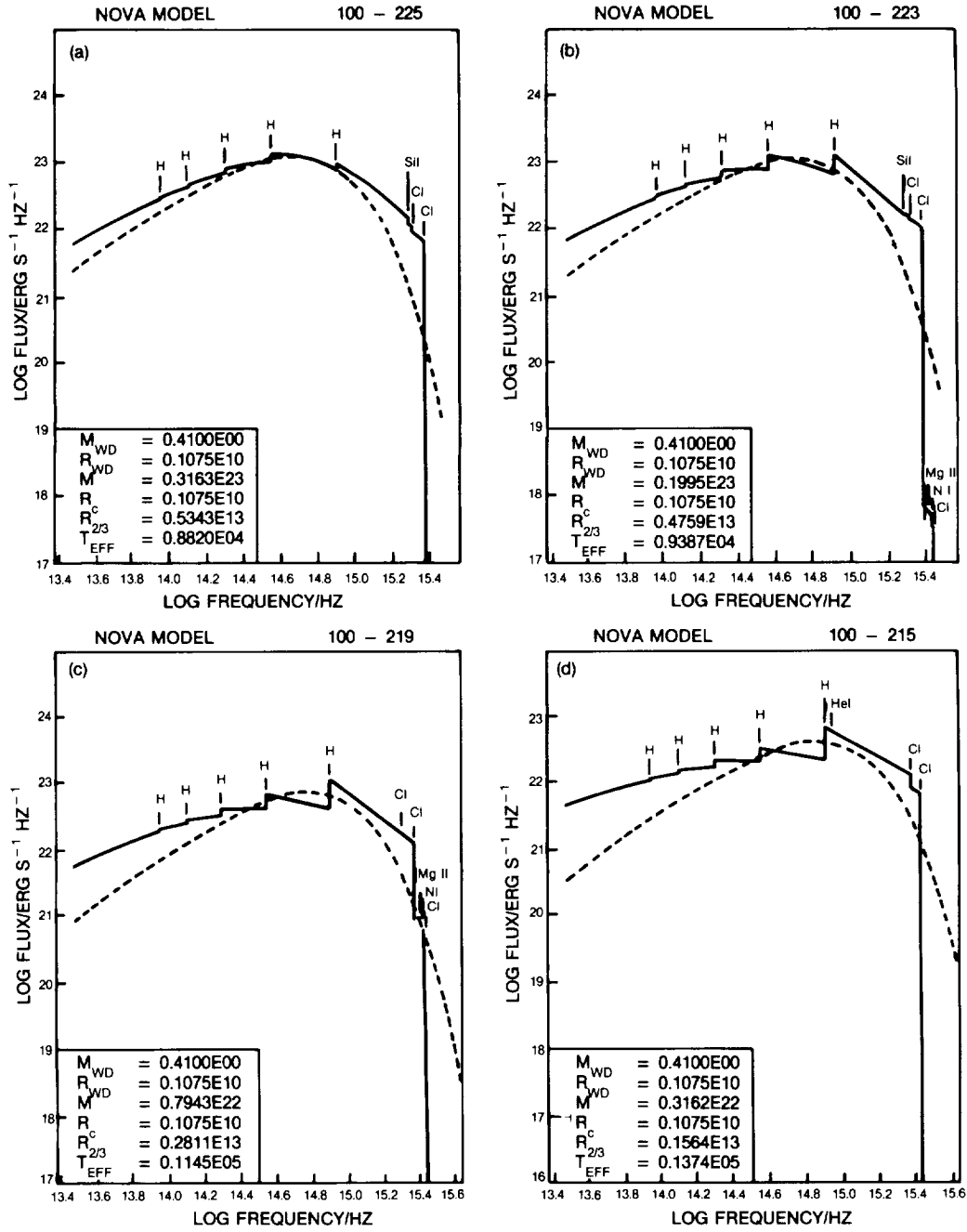


Figure 7-7. Emergent flux distributions for steady-state nova winds with (a) $\log \dot{m} = 22.5$, (b) $\log \dot{m} = 22.3$, (c) $\log \dot{m} = 21.9$ and (d) $\log \dot{m} = 21.5 \text{ g s}^{-1}$ (Harkness, 1983).

III. D. CONTINUED EJECTION B

A form of continued ejection is possible, where most luminosity arises from previously ejected material, ejected at a time when the mass-loss rate was very high. Such a model, therefore, resembles instantaneous ejection Type I with an envelope of constant thickness, if the ejection velocity is constant at the time when most of the ejection takes place. The observational consequences are very similar.

Continued ejection B can be expected to occur when the mass-loss rate decreases rapidly. A necessary (but not sufficient) condition for continued ejection B can be derived from Friedjung (1966a), if the mass-loss rate \dot{m} varies as the power of ejection time t_0 from optical maximum (except of course for times close to the latter), then

$$\dot{m} = \dot{m}_0 t_0^{-\alpha} \quad (7.12)$$

In the case when continuum emission in optically thin regions is due to recombinations at constant temperature and is proportional to the density squared, the necessary condition for continued ejection B to be valid for continuum radiation is

$$\left(\frac{R_p}{R_i} \right) \left(\frac{R_0}{R_0 - R_i} \right)^{2\alpha} > 1. \quad (7.13)$$

Here R_i is either the outer radius of the part of the envelope where hydrogen is completely ionized, or the radius at which the power law of Equation (7.12) breaks down when the latter is smaller. R_0 is the outer radius of the envelope. When $R_0 - R_i > R_p$, condition (7.12) requires $\alpha > 0.5$. In addition, condition (7.13) suggests that continued ejection B is unlikely unless the hydrogen in the envelope is ionized almost to the outer edge.

The conditions for line emission need not be the same as for continuum emission. Strong lines, including in particular Balmer lines, though formed by recombination, may be optically very thick. Photons scattered many times in optically very thick lines can be lost through

other processes. In view of such effects, line emission from outer parts of the envelope might dominate, even in situations where the continuum from such regions is relatively weak.

III.E. CENTRAL STAR DOMINANT MODELS

Ejection is supposed to occur from one of the components of the central binary, and one can imagine a general swelling of one of the components, so that something resembling a normal, almost stationary, stellar photosphere is observed after optical maximum. In this case, the high velocity expanding layers whose presence is deduced from the spectrum must be optically thin in the continuum. This type of model, which played a role in the history of nova models, more recently became attractive to those theoreticians who do not examine observational constraints in great detail.

A summary of older work on central star dominant models was given by Mustel, (1957). Figure 7.8 is taken from his work to illustrate the model. A fundamental change was supposed to take place at optical maximum with the detachment of the outer parts of an ex-

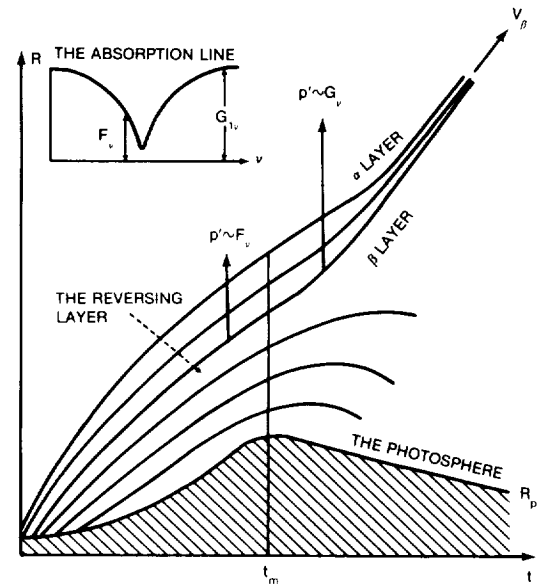


Figure 7-8. A representation of Mustel's (1957) model, in which the outer parts of the extended reversing layer of the nova continue to rise in the form of an envelope, while the star contracts.

tended reversing layer. Mustel considered the reason for the detachment to be a different radiation pressure, though in a later paper (Mustel 1962), cosmic ray pressure was proposed. In this model, the shell sweeps up preexisting premaximum system material, the shell itself being where the principal system is formed. The star itself then contracts after optical maximum.

Theoretical work to be described later suggests that the white dwarf component of the underlying binary expands during an outburst. Such a white dwarf can reach giant dimensions after having engulfed its companion, and its photosphere could give rise to the continuum emission. Any continued ejection would resemble the winds of "normal" hot stars, which are optically thin in the visual continuum. Indeed, in view of the large luminosity of a nova for a long time after optical maximum, strong winds of this kind are expected to be present.

Central star dominant models run into a major difficulty because of the usual lack of unshifted absorption lines and of relatively narrow unshifted emission lines from the region where the wind might be accelerated. One might expect to most easily observe photospheric absorption lines at times and for lines for which the line emission due to the expanding layers is relatively weak. One attempt to detect the strong unshifted absorption lines expected for FH Ser, if it had the photosphere of a normal giant with the same temperature and luminosity, indicated that such lines were not present (Friedjung (1977b).

In the framework of a central-star-dominant model, as well as for continued ejection A, the change in the energy distribution of a nova as it fades needs to be interpreted by an increase in photospheric temperature. In fact, lines of more highly ionized states are seen later. If a central-star-dominated model were true, one expects to see the lines of a high state of ionization formed in the photosphere (unshifted absorption lines) or in a chromosphere (narrow emission lines), at stages when no contribution to the line profiles from expanding material is

seen. In fact, such a situation is never observed, and, as we shall discuss later (and have discussed before), the most central regions have the highest velocities.

According to McLaughlin (1943), faint non-displaced absorption lines briefly existed around maximum light in the spectra of DN Gem, GK Per, and perhaps also V603 Aql. McLaughlin considered such lines circumstellar (seen because of radiative excitation of their lower levels near nova maximum), and in any case, one might expect their appearance in a much more systematic way, if they were the photospheric absorption lines of central-star-dominated models. Narrow line emission was seen in the spectrum of HR Del *before* optical maximum. One can wonder whether during this stage, not seen for other well-observed novae, a central-star-dominant model may not be the best way of describing the situation.

If we summarize the discussion of these simple models, instantaneous ejection Type II and central-star-dominant models put strong constraints on the velocity distributions; according to them, the lowest velocities must be near the centre of the envelope. The velocity distributions of such models encounter observational objections, which will be mentioned in more detail later. Instantaneous ejection Type I would require the continuum to be formed either in a thin shell or in a "nova remnant"; blueshifted P Cyg absorption components should be very wide in the first case and very narrow in the second one. In the first case, one would not expect to see absorption of the continuum by expanding material inside the shell; as will be seen, major difficulties are encountered for this type of model. Continued ejection A has also strong constraints; the continuum brightness is directly related to the mass-loss rate, and, if the inner regions have higher velocities than the outer ones, collisions between faster and slower moving material should accelerate the later following momentum transfer.

III.F. COMBINATION OF SIMPLE MODELS

The foregoing discussion suggests that the most fruitful way of making progress is to combine the approaches of continued ejection A and the presence of a thin shell. In many cases, at least, most radiation of the continuous spectrum is understood most easily as coming from an optically thick wind, while most line radiation may, at least sometimes rather come from a shell. Other reasons and considerations will be given later.

It can be noted that it is possible to combine continued ejection A with a central-star-dominant model, if electron scattering is much larger than pure continuous absorption. In such a case, P Cyg lines would only be seen clearly for those layer that are optically thin to electron scattering, while the continuous spectrum would come from deeper layers, which would have a lower velocity. Such a situation was proposed by Turolla et al. (1988) to occur in Wolf-Rayet stars. Presently, reasons will be given against this type of interpretation in the case of novae, but its possibility should be kept in mind in future studies.

It may be useful in this connection to emphasize the differences between novae and Wolf-Rayet stars, which like novae in outburst, also appear to have a very large mass-loss rate. In particular, Wolf-Rayet stars show signs of a velocity gradient in their winds. The shifts of the violet-displaced absorption components are correlated with line excitation potential but not so simply related to the ionization potential. This can be understood if the lines are formed in an accelerated wind (Willis and Garmany, 1987). A tendency is thought to exist for ionization to be frozen in such winds, explaining the contradictory results when one attempts to correlate velocity with ionization potential. As shall be seen, the picture that emerges for novae is rather different, though the lines of the Orion absorption system need to be studied in more detail to completely eliminate the possibility of such effects for them.

The higher velocities of later ejected material observed near the centre of the envelope suggest another model (Friedjung, 1987b), which can be considered a theoretical development of the consequences of continued ejection A, to be compared with the observations. A high-velocity wind is supposed to interact with slower moving material ejected before optical maximum, the latter being expected to produce the premaximum absorption system. A thin shell is formed by a snowplough effect associated with the collision of the two regions. This shell is assumed to be the seat of the principal absorption system and its associated emission, and when the postoptical maximum activity of a nova ceases, this shell can be expected to contain most of the ejected mass.

The model described is in many ways a combination of the simple models described previously. The premaximum system material, ejected before optical maximum and then swept up, might be partly described at least by instantaneous ejection Type II, the shell by instantaneous ejection Type I, and the wind by continued ejection A.

To predict the consequences of the model, the theory of the interaction of a stellar wind with the interstellar medium, described, for instance, in the review of McCray (1983), can be used. In early stages, densities are high, and the shocked material cools rapidly, directly transmitting momentum to the shell. In later stages, the time scale for cooling becomes longer than the characteristic time for the density of the wind near the shell to decrease because of the expansion of the latter. Shocked plasma then stays hot, exerting a thermal pressure on the shell, and also tends to fill a large proportion of the volume inside the shell. The condition for the transition between the two situations is estimated using order of magnitude arguments by Friedjung (1987b) as

$$t_a \geq \frac{8.0 \times 10^{-21} T^{-1/2}}{m_\mu^2} \frac{\bar{m}}{V_p^3 (V_w - V_p)^2 f}, \quad (7.14)$$

with \bar{m} being a mean mass-loss rate for the wind; T , the hot plasma temperature; m_μ , a mean mass for the atoms and ions present; V_w ,

the wind velocity; and V_p , the shell velocity. t_a is the time since ejection when the transition takes place, the hot plasma then filling a fraction f of the volume inside the shell. In this expression, the formula of Kahn, (1976), which approximates the cooling rate of Raymond et al. (1976) was used. Taking $f = 0.3$, and $V_p = V_w - V_p = 1 \cdot 10^8 \text{ cm s}^{-1}$, t_a is found to be $5 \cdot 10^5 \text{ s}$. In this calculation, $\dot{m} = 10^{22} \text{ g s}^{-1}$, $T = 3 \cdot 10^7 \text{ K}$ and $m_p = 3 \cdot 10^{-24} \text{ g}$. Such a result however, is sensitive to both the numerical values and the physical assumptions used.

A model of this kind leads to a number of predictions. The shell should be accelerated by the pressure exerted on it. In addition, one can explain why the velocity of the premaximum system appears to increase after optical maximum. For instance, if the ejection of the material of this system takes place according to instantaneous ejection Type II, the fastest material is at the outside and is swept up last. One can precisely calculate when material, having the velocity of the last seen premaximum system, should have been ejected, assuming that it suffered no acceleration. The result can be compared with the time when material having this velocity was first seen before optical maximum for the few novae well observed in this stage. Finally, the shocked plasma should emit X ray and coronal emission lines: the former might be expected to be strongly absorbed by the shell in earlier stages. Calculations of the predicted X-ray and coronal line emission that should be observable indicated that they might be rather weak, being so masked by emission produced by other processes.

Other models involving collisions have also been previously proposed. Bychkova and Bychkov (1976) considered the collision between principal and premaximum system material, with the production of inward and outward propagating shocks. Bychkova (1982) later considered collisions between continuously ejected material and that of the principal system. Both regions were supposed to be extremely inhomogeneous, so collisions occurred between fast and slow-moving blobs. The diffuse-enhanced and Orion absorptions were supposed to be produced in shocked plasma,

while most of the radiation of the continuous spectrum was also supposed to be produced by the colliding material. Therefore, the observed phenomena would be due to processes near the outer edge of the envelope. It would then be difficult to explain rapid variations such as those observed for DK Lac several months after maximum, over time scales of the order of 10^{-2} of the time elapsed since maximum. In addition, the author of these lines has difficulty in understanding how an apparently optically thick continuum energy distribution for emitted radiation could then be produced for novae.

In any case spherically symmetric collision models are probably too simple. The observed deviations from spherical symmetry need to be taken into account. An early attempt to do this was made by Hutchings (1972). Certain velocities were supposed to occur only in certain directions in order to explain the observed emission line profiles. Slow-moving principal system material, after interaction with the companion star, might only occur in polar cones. It should be noted, however, that the principal absorption system is always seen in the spectra of classical novae; this suggests that the material associated with it is present in all directions around a nova. The densities etc. are, of course, presumably dependent on the directions, and the orientation of the structure of the envelope, with respect to the observer, needs to be taken into account much more than in the past.

III.G. RECURRENT NOVA MODELS

The observed characteristics of recurrent novae in general, and of RS Oph and T CrB in particular, are rather different from those of classical novae; thus, models for these stars need not be the same. Let us recall that, in the course of an outburst, the observed expansion velocity of the absorption lines and the associated wide emission lines of RS Oph and T CrB decrease with time. Narrow emission components seen in the spectrum of RS Oph disappear during the same development.

These observations led Pottasch (1967) to propose a different kind of model for RS Oph.

An ejected shell was slowed down by a pre-existing circumstellar envelope, the latter being where the narrow emission lines were formed. He made quantitative estimates of the envelope mass assuming the same type of increasing envelope thickness as for a classical nova (the electron temperature and the shell thickness increase were assumed to be slightly higher). Knowing the thickness and the radius of the shell, hydrogen and helium fluxes gave the total mass, following theoretical expressions for the emission line intensities. Assuming momentum conservation, the density distribution of the preexisting circumstellar envelope could be determined. Pottasch found that it appeared to be in hydrostatic equilibrium, with a density distribution characteristic of a temperature of 10^4 K in inner regions and 10^3 K in outer regions. The resulting mass of the central star was $0.7 M_{\odot}$.

T CrB and RS Oph are at present thought to be closely related to symbiotic stars, which are now considered to be binaries consisting of a cool giant and a more compact companion. Cool giants have strong winds, and the material of this wind would be swept up by the ejected material of a nova explosion. The basic idea of Pottasch's (1967) model is therefore attractive, even though details may be considerably different.

Pottasch's ideas were further developed by Gorbatskii (1972, 1973). The formation of a shock in the circumstellar envelope ahead of the shell was considered. Coronal line emission was studied, taking into account the different time variation of electron and ion temperature of the model. Gorbatskii (1977) suggested that similar ideas could also explain the narrowing of emission lines in the classical novae V1500 Cyg and CP Pup; red giants, however, are not present in these objects, and instantaneous ejection Type II may best describe the development of a large part of the envelopes of these exceptionally fast classical novae. Hydrodynamic calculations of what happens when high velocity ejection takes place in a low velocity wind were performed by Bode and Kahn (1985). They considered a model similar to that

of supernova remnants: A spherical envelope expands into the wind of a cool giant, and forward and backward shocks are generated. Matter that has passed through the reverse shock is well cooled, forming small condensations following a Rayleigh-Taylor instability. In what the authors call Phase II, the movement of the backward shock confines unshocked gas to a progressively smaller volume in the middle, while, at the outside, a blast wave advances into the wind. The newly shocked outside gas is very hot and does not radiate much of its energy. In Phase III, cooling of the outer shock dominates and condensations are again formed, which can break loose and move at high velocity in the wind. The model was applied to explain X-ray and non-thermal radio emission.

The authors calculated that the first X-ray observations were made at a transition between Phase II and Phase III, the temperature from the shock being between 0 and 5.7×10^6 K. Analysis of the X-ray observations seemed to require a "metal" overabundance of about 5 in the *red giant wind*. The authors also concluded that the presence of optical emission lines throughout the development of RS Oph requires parts of the shocked gas to cool earlier; they supposed that denser condensations were produced by a magnetic field. This giant wind magnetic field was calculated to be 0.01 G ahead of the shock, and 0.04 G in the shocked gas; relativistic electrons moving in the field would then have produced the radio emission non-thermally.

Even if this type of model stands the tests of more observations and can be developed further, it probably cannot be applied to all recurrent novae. Not all of them appear to have cool giant companions; however, it is not clear to what extent classical nova models can be applied to recurrent novae without giants.

IV. EMPIRICAL APPROACH

At this point, we need to see whether observational results can be pushed further, in order to give a more precise indication of what is happening. Without already having a detailed

model, is it possible to do basic diagnosis? This type of approach will now be described.

IV.A. VELOCITY STRATIFICATION FOR CLASSICAL NOVAE

The different absorption components observed, which can sometimes be very numerous, suggest that motions are not simple. However, observations do suggest a definite stratification, often not taken into account in models.

Evidence from the study of the optical lines was summarized by McLaughlin (1947). He stated that in practically all detailed studies of classical novae, four absorption systems can be recognized, which are the pre-maximum, the principal, the diffuse-enhanced, and the Orion systems. These systems can be split into sub-systems, so that in certain situations the total number is much greater, but this does not invalidate McLaughlin's classification. In any case, this classification is both chronological and in order of velocity; the premaximum system appears first with the lowest velocity and the Orion system last with the highest velocity, already indicating that low-velocity material is ejected first and so is further from the ejecting star at a given time than high-velocity material ejected at a later time.

McLaughlin (1947) was guided by three main considerations: (1) superposition of line profiles of different systems, (2) response to disturbances originating in the ejecting star, and (3) excitation and other physical processes. Considerations of the first type could be applied when diffuse-enhanced or Orion components of line profiles were superposed on lower velocity components. Principal system absorption components were not filled in by diffuse-enhanced or Orion emission from the profiles of other lines: a striking example was that of DQ Her, where the principal absorptions of Sc II 4247 Å remained strong and sharp when superposed on the longward wing of the Fe II 4233 Å diffuse-enhanced emission; similarly, Orion N III emission did not fill in principal and diffuse-enhanced absorption. On the other

hand, diffuse-enhanced absorption lines were often partially or completely obliterated by overlying principal emission. McLaughlin quotes examples for DQ Her. Cases of Orion absorption being disturbed by overlying principal emission were also quoted; in the case of V603 Aql, each Orion absorption component of the N III pair near H δ weakened in turn as it coincided with the maximum of O II principal system emission. All this clearly suggests that higher velocity material is *below* that of lower velocity.

The other considerations of McLaughlin (1947) also pointed to the same conclusion. When secondary oscillations of light occur, the Orion absorptions, unlike other absorptions at such stages, show close wavelength correlations with brightness changes; as will be seen, this is also true if the brightness of the continuous spectrum is considered. In the framework of a continued ejection A model, this would suggest a correlation of a Orion system velocity and the ejection rate. Finally, the higher ionization of the Orion system suggests an origin in the inner parts of the envelope, if photoionization by radiation from the central photosphere dominates. McLaughlin explained the disappearance of high ionization bands in the principal spectrum during "flaring" of the 4640 Å and other bands, supposing complete absorption of high-frequency radiation by the inner envelope during such stages.

As a careful experienced analyst of optical spectra, McLaughlin's arguments carry great weight, and his conclusions are very probably correct. However, future studies of line superpositions need to be quantitative. The interpretation is not always obvious. In addition, it may also be noted that emission not only fills in absorption lines produced in deeper layers, but can also fill in absorption lines in another line of sight. Large deviations from spherical symmetry are necessary for this to be important.

Since the classical work of McLaughlin, very high-velocity absorption components (up to 10^4 km s $^{-1}$) have been detected in the satellite ultraviolet spectra of some novae. It is not yet

completely certain whether they can be fitted into the classification for optical spectra. However, the highest velocity systems of V1370 Aql varied on a time scale of a few hours, suggesting line formation in the inner envelope.

Other points concerning velocity stratification need to be made. Absorption components of the principal absorption system are always seen for classical novae; this means that deviations from spherical symmetry do not appear to be large enough to prevent the formation of the system in any direction. High-velocity systems are also generally seen; there has been some uncertainty in the case of V1500 Cyg; however, even in the case of this exceptionally fast nova, Duerbeck and Wolf (1977) identified the presence of diffuse-enhanced absorption about one day after maximum. Therefore, limits are also placed to possible deviations from spherical symmetry for the high-velocity systems.

If the velocity identification of McLaughlin is accepted, this places strong constraints on the region of production of the continuous spectrum. A large part of the continuous spectrum *at least* must be produced below the level where any strong absorption line is formed. Therefore, when strong Orion N III absorption components are seen, at least a large part of the continuum must be produced in more inner regions, a condition compatible with continued ejection A or central-star-dominant models. Such was the case for V603 Aql (Friedjung 1968). However, extrapolations of such conclusions to other epochs of nova development carry some uncertainty; as will be seen below, the best way is to reason from regularities in the time variation of the continuum flux.

IV.B. ANALYSIS OF THE REGIONS WHERE THE CONTINUOUS SPECTRUM IS PRODUCED

In continued ejection A and central star dominant models, most of the continuum is emitted by a photosphere, although in the former case, one has a "quasi-photosphere" formed by an optically thick wind. However, it

should be noted that infrared emission is not included in such considerations; the outer parts of the envelope may be expected to be optically much thicker in the infrared than in the optical because of free-free (and sometimes dust) opacity.

The basic ideas of analysis have already been given in Section III.C. Photospheric temperatures and radii can be found, which, when continued ejection A is assumed, can be converted into mass-loss rates. However, all calculations made up to now are extremely approximate and of dubious physical consistency. Among temperature determinations, color temperatures are probably the easiest to interpret; other methods such as Zanstra-type temperatures do not only make assumptions about the relation of the energy distribution of emitted radiation to the photospheric effective temperature, but also about the excitation of emission lines. The latter temperatures assume production of line emission by photoionization, followed by recombinations and cascades to the ground state, all ionizing photons being absorbed by the line-emitting medium.

In view of this, we shall emphasize analysis, which is close to observational data and which should be easily reinterpretable in the future using better theory. The first step is to see what optical continuum fluxes (expressed as magnitudes) can tell us. When continuum magnitudes are plotted against log time from maximum, the graphs obtained have often linear portions; this means that flux varies as a power of the time from maximum (or perhaps rather from the initial explosion, which occurred not much earlier). Such graphs are shown in Figure 7.9, 7.10 and 7.11. The first two of these are for V603 Aql and GK Per, which showed oscillations during their declines; these oscillations were between parallel lines associated with early and late decline. The difference between the continuum flux magnitude and the visual magnitude, which includes the effects of emission lines, can be seen by comparing Figures 7.9 and 7.12, the lines not being parallel in the latter figure. The linearity of such graphs for visual magnitudes was shown by Vorontsov-Velyaminov (1940).

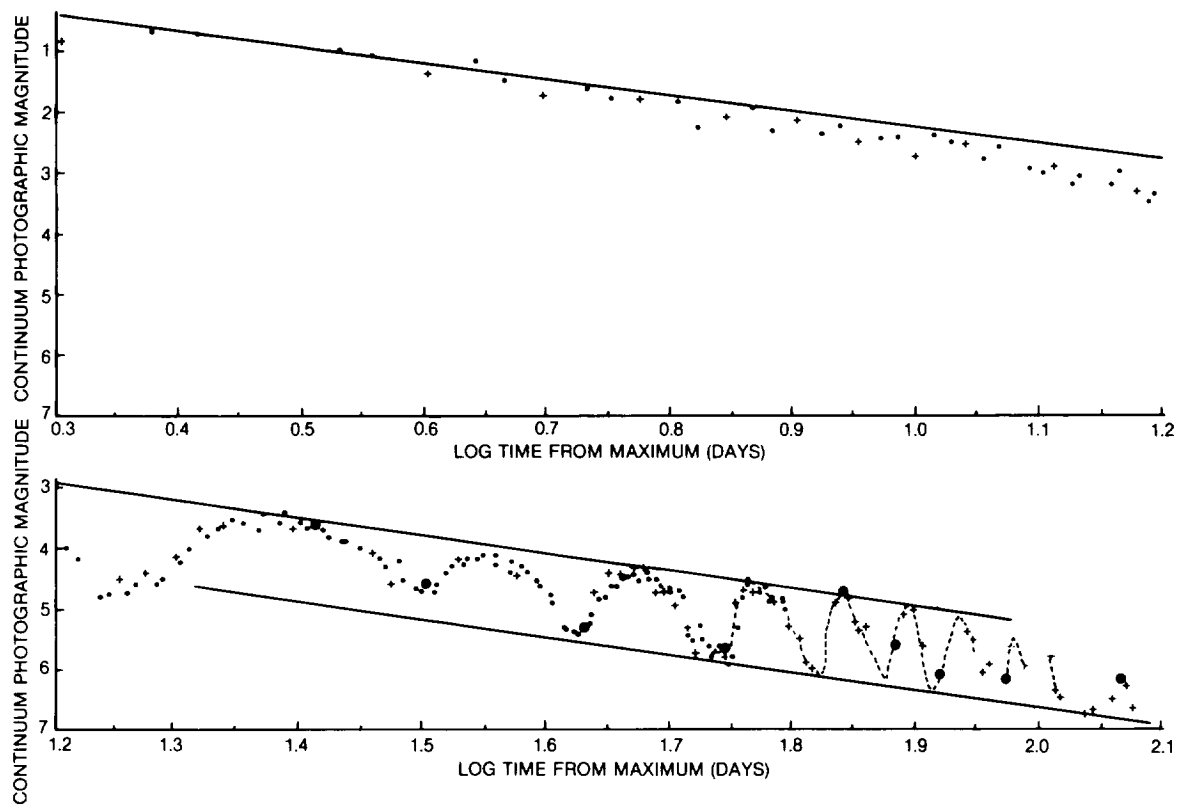


Figure 7-9. The continuum magnitude of V603 Aql versus log time from maximum (Friedjung, 1966a).

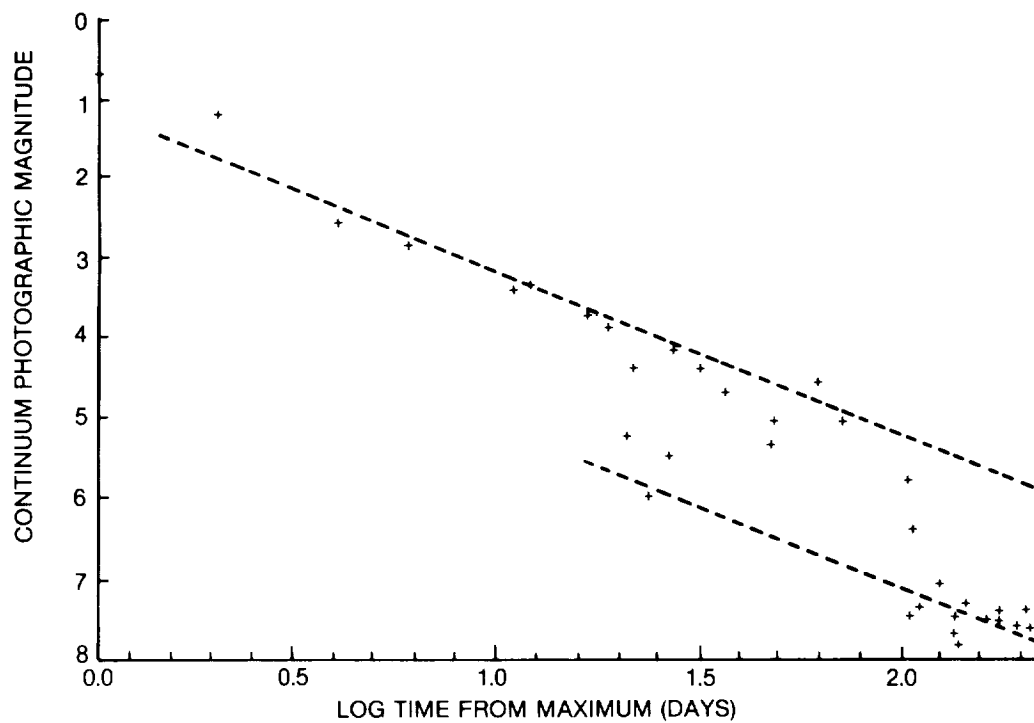


Figure 7-10. The continuum magnitude of GK Per versus log time from maximum (Friedjung, 1966a).

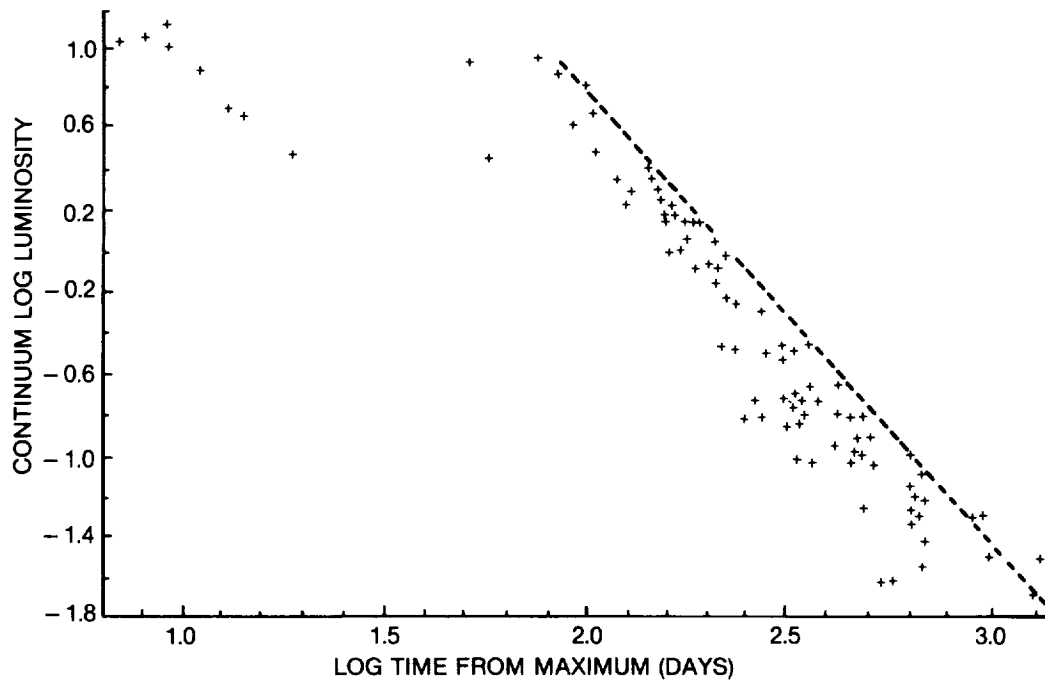


Figure 7-11. The continuum magnitude of RR Pic versus log time from maximum (Friedjung, 1966a).

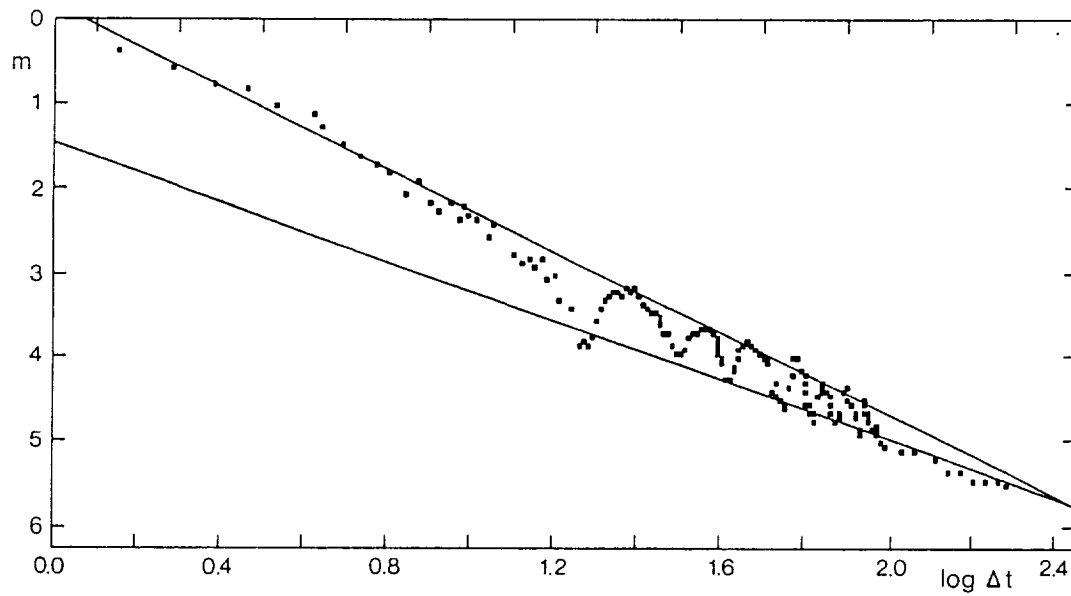


Figure 7-12. The visual magnitude of V603 Aql versus log time from maximum. The behaviour of the oscillation can be clearly seen (Friedjung, unpublished).

For novae such as V603 Aql and GK Per, one can conclude that when, during the decline, the continuum magnitude varies as the same power of the time, the basic physics and hence the most suitable model very probably do not change. In addition, if oscillations occur between parallel lines, as in that graphs described in the last paragraph, it is tempting to conclude that similar physical processes occur at both maxima and minima. However, the last conclusion is much less certain.

Continuum magnitudes combined with Zanstra temperatures were used by Friedjung (1966b) to derive radii, assuming a Planckian energy distribution for the photosphere. These radii appeared to have the same type of power law variation as the continuum fluxes. In spite of the doubts that can be cast on such a calculation, it may be that the conclusion concerning the power-law time variation is not strongly dependent on the assumptions. Such a hypothesis needs obviously to be confirmed. If the radii follow a power-law time variation, there is moreover a good chance that the same is true for the mass-loss rate.

In view of the lack of a reliable theoretical model to give the whole distribution of energy emitted by a nova, the observations in other spectral regions are needed for the determination of basic data, e.g., those concerning the total luminosity. The combination of observations in different spectral regions shows that the total luminosity declines much more slowly than in the optical, and indeed may stay almost constant for a long time. Gallagher and Code (1974) studied the time variation of the radiation from FH Ser between 1550 and 5480 Å, and found it almost constant for more than a month after optical maximum (Figure 7.13). However, if one attempts to correct for emission in other spectral regions indications of a decline are seen (Friedjung 1977b) while, in any case, conclusions are sensitive to the reddening corrections. More recent results for other novae are shown in Figures 6-46 and 6-50. The slowness of the decline of integrated flux is obvious.

The total luminosity of novae was also stud-

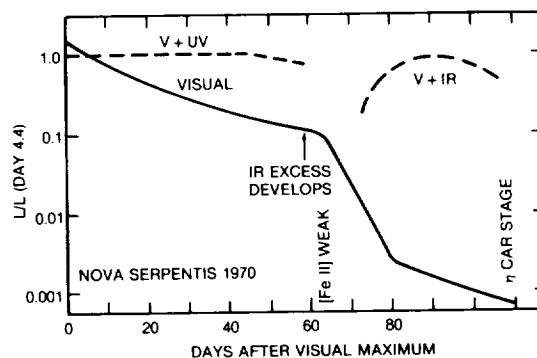


Figure 7-13. Energy budget of FH Ser as a function of time, based on ultraviolet, optical, and infrared data. For a distance of 650 pc, the day 4.4 luminosity is $1.6 \times 10^4 L_{\odot}$. The postmaximum luminosity plateau, the correlation between optical, spectral, and light curve features and the development of the thermal infrared excess are shown (Gallagher, 1977).

ied by Duerbeck (1980), using ground-based data. Absolute magnitudes were determined from newly found distances and interstellar extinctions. A bolometric correction corresponding to a mean spectral type of F5Ia (-0.25) at optical maximum was used to obtain the luminosity at optical maximum. Duerbeck found that fast or moderately fast novae with smooth declines (except for transition-stage oscillations in some cases) had luminosities well above the Eddington limit, while slower novae had luminosities in the region of the Eddington luminosity. In any case, this type of calculation is still extremely approximate.

Novae for which multifrequency observations are available can be studied further. One can not only attempt to determine the total radiative luminosity, but also the luminosity associated with the kinetic energy of the wind, which is large for continued ejection A. Such an attempt was made by Friedjung (1987a) for FH Ser, using rather approximate theory. The energy distribution had been studied from the infrared to the ultraviolet, and an examination of observed energy distribution indicated that a blackbody fit was not too bad, thus enabling a photospheric color temperature to be defined. A corresponding blackbody photospheric radius R_p could then be derived. If the optical depth in the photosphere is assumed to be 2/3,

the kinetic energy flux is

$$F_K = \frac{4\pi R_p V_p^3}{3\kappa}, \quad (7.15)$$

with V_p the ejection velocity, and κ the opacity supposed to be dominated by electron scattering. V_p was taken to be the higher of the observed absorption component velocities, because, as seen above, higher velocity material appears to be closer to the photosphere. Friedjung (1987a) obtained what are probably rather minimum values of F_K , as the measured V_p corresponding to the mean absorption component radial velocity was an average of ejection velocity components in the direction of the observer, while a maximum κ of 0.15, suggested by the calculation of Bath (1978), was taken. The total radiative and kinetic energy luminosity found and shown in Table 7-2 appears to remain for a long time above the Eddington limit of $2.07 \cdot 10^{38} \text{ erg s}^{-1}$ for a $1 M_\odot$ star with a chemical composition characteristic of a nova as given by Stickland et al. (1981). This result however, is approximate in view of the assumptions mentioned, while it is also clear that the maximum κ is sensitive to the various element abundances. Therefore, this type of calculation needs to be repeated with better theory in the future.

Other conclusion that can be drawn from Table 7-2 should also be emphasized. The kinetic energy flux is of the same order as, and indeed somewhat larger than, the radiative flux. In view of the fact that the velocity of the continuously ejected wind appears to be of the order of 0.005 the velocity of light, the ratio of the momenta of radiation emitted in unit time to that of material ejected in unit time is of the order of 3×10^{-3} . The ratio of the radiative energy per unit volume to the kinetic energy per unit volume is not much larger near the photosphere. This suggests that, unless the estimates of Table 7-2 are wildly wrong, acceleration of the wind by radiation pressure to the observed velocities cannot be produced at small optical depths. Acceleration by radiation pressure in the lines appears to be quite insufficient. However, radiation pressure can act in another way. Equations 7-5 and 7-6 indicate that the ratio of energy densities can be much larger at

large optical depths; when electron scattering dominates, Equation 7.6 leads to a variation as r^{-1} for a constant velocity, so the ratio would be of order unity at 10^{-3} to 10^{-2} of the photospheric radius.

It is at such radii that radiation pressure might be responsible for accelerating the continuously ejected material. It is for this type of reason, that the combination of continued ejection A and the formation of the continuous spectrum in a low-velocity photosphere, described above, is hard to reconcile with acceleration by radiation pressure.

Similar conclusions were previously reached by Friedjung (1966c), these being, however, based on temperatures and radii deduced only from ground based observations. The improved multifrequency photospheric temperatures and radii have not changed the situation radically, and one might perhaps doubt whether better diagnostics could really make such a large difference, in spite of the present approximations.

The velocity variations of the absorption lines of the Orion system can be closely related to the behaviour of the continuous spectrum. Very often there seems to be a correlation between the velocity of the Orion system and the brightness of the continuum. Correlations of velocity squared with $1/\text{radius}$ from Friedjung (1966c) are shown in Figures 7-14, 7-15, and 7-16. The radii of V603 Aql and RR Pic derived from Zanstra temperatures assuming a black-body energy distribution and those from color temperatures of DQ Her show almost linear correlations. What is also very striking is that the velocity at infinite radius corresponds to that of the diffuse-enhanced system. This suggests that both absorption systems are due to the same physical process, best understood as continued ejection.

It may be noted that GK Per, another nova showing like the previous ones postoptical maximum oscillations, did not appear to have a velocity-radius correlation, according to Friedjung (1966c). It now seems (Bianchini et al., 1988) that its velocity at a given time has oscillations with twice the instantaneous period of the postmaximum light oscillations.

TABLE 7.2 PHOTOSPHERIC PROPERTIES OF FH SER

day from 1970 Feb. 14.0	L in 10 ³⁸ [erg s ⁻¹]	T _c [K]	R _p 10 ¹² [cm]	mean measured high vel. low vel. [10 ² km s ⁻¹]		F _k (min) [10 ³⁸ erg s ⁻¹]	F _k (min)+L
6.39	2.65	5250	22.0	most hydrogen neutral			
8.41	2.39	5370	20.1	—	—	—	—
15.85	1.32	7410	7.8	13.1	6.7	4.9	6.2
22.05	1.05	9120	4.6	15.1	7.1	4.4	5.5
27.34	0.92	9770	3.7	16.3	7.3	4.5	5.4
29.34	1.39	8320:	6.3:	16.8	7.4	8.3:	9.7:
31.87	0.91	9200	4.3	17.2	7.4	6.1	7.0
49.83	0.73:	14800:	1.5	18.5	7.6	2.7:	3.4:
57.49	0.64	18600:	0.84:	18.7	7.7	1.5:	2.2:

L = radiative luminosity; T_c = color temperature, R_p = photospheric radius; vel. = (expansion) velocity; F_k = kinetic energy flux.

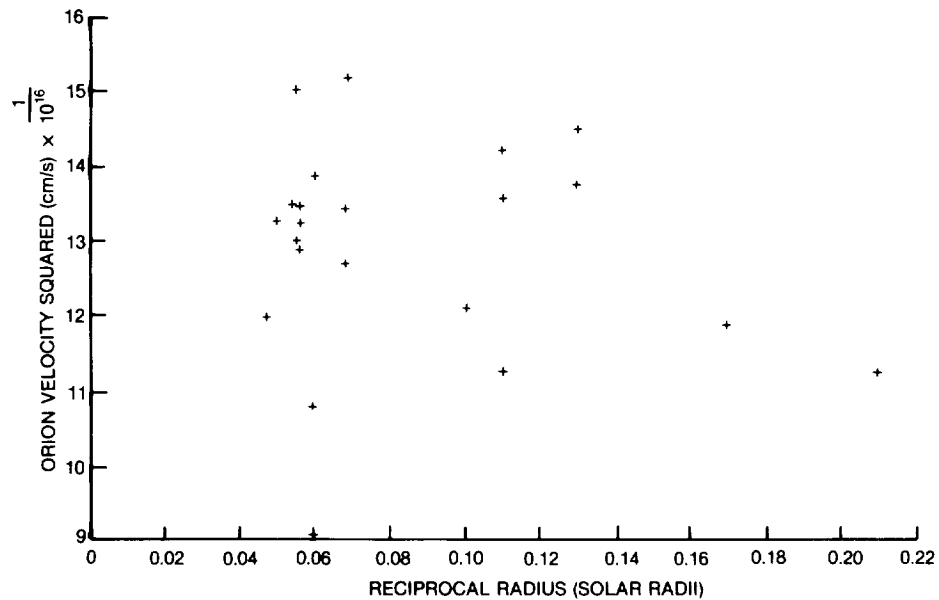


Figure 7-14. The relation between velocity squared and reciprocal radius for GK Per (Friedjung, 1966c).

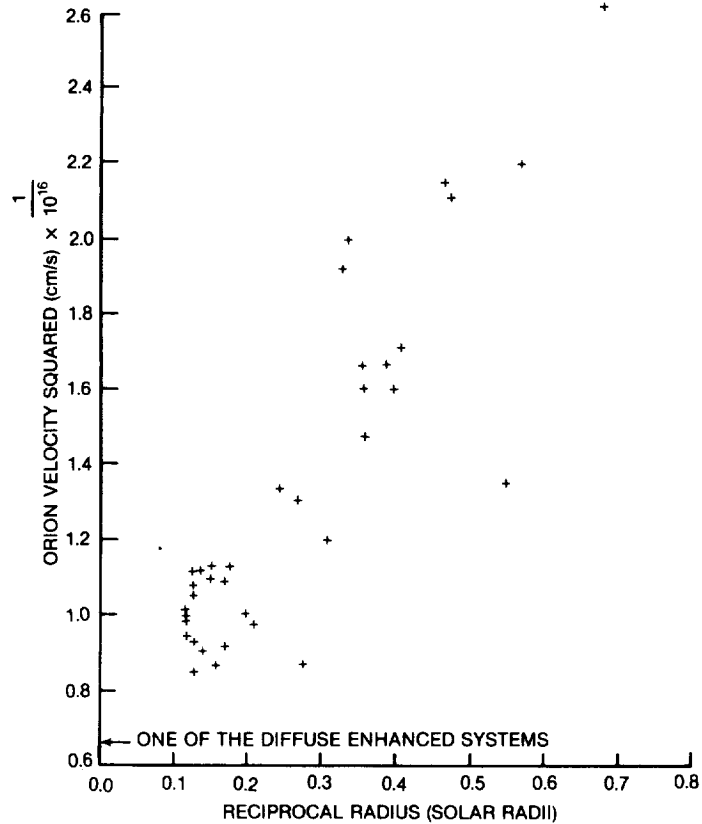


Figure 7-15. The relation between velocity squared and reciprocal radius for RR Pic (Friedjung 1966c).

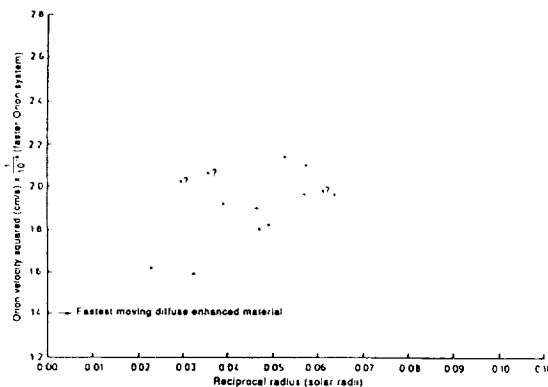


Figure 7-16. The relation between velocity squared and reciprocal radius for DQ Her (Friedjung 1966c).

The most attractive picture that emerges from these considerations seems to be one of wind acceleration by radiation pressure at large optical depths, where the total luminosity, at least in a limited region, is above the Eddington limit. Part of this total luminosity is converted to kinetic energy. It is clear that such a process,

if it exists, cannot be thermal! Its possibility will be discussed later.

IV. C. ELEMENT ABUNDANCES IN NOVA EJECTA: CURVE OF GROWTH METHOD

Study of emission and absorption lines in principle, can give information about abundances. Sometimes, extremely abnormal abundances have been determined. Nevertheless, such determinations have traps, which should not be neglected.

Two types of method can be considered. The first uses absorption components of lines, and applies curve of growth methods to derive abundances. The second uses emission lines, which, particularly in the nebular stage, should be formed under conditions similar to those of planetary nebulae, for which one knows how to determine abundances. Older work is summarized by Collin-Souffrin (1977) and by Williams (1977). The subject has expanded very

much since these reviews.

In the analyses of Mustel and his coworkers, using absorption lines and the curve of growth method, nova spectra were compared with those of stars with "similar spectral classes" (F supergiants). The narrowness of the observed nova absorption components suggested that they could be treated in the same way as the lines of normal stars, in spite of the blueshift due to the expansion (in fact, such narrowness could be produced when absorption lines are formed in an envelope where most line absorption is at a radius much larger than that of the photosphere). A systematic introduction to the curve of growth method can be found in Mustel (1964). The partial curves of growth are constructed, using equivalent widths of certain ions, are shifted in abscissa to make them coincide, and then are compared with theoretical curves of growth. A problem is encountered with the excitation temperature, T_{exc} . The diagram showing multiplet strength versus excitation potential does not result in a straight line, whose slope is given by T_{exc} , but shows an overexcitation of levels with high excitation potential, which might be explained by a temperature variation in the extended layer, or by isolated high-temperature cells in the extended atmosphere.

The curve of growth method was applied to the novae DQ Her, HR Del, and V1500 Cyg. Mustel and Boyarchuk (1959), Mustel and Baranova (1965), Antipova (1974) and Mustel (1974) based their analysis on the premaximum system of DQ Her, Mustel and Baranova (1966) analyzed the principal absorption system of this nova one week after optical maximum. HR Del was analyzed by Ruusalepp and Luud (1971), Antipova (1974) and Yamashita (1975). Premaximum and maximum spectra of the fast nova V1500 Cyg were analyzed by Boyarchuk et al. (1977). Another approach, based on simple synthetic premaximum spectra assuming LTE, was carried out by Stickland (1983). Generally, overabundances in C, N and O were found [though Stickland (1983) found C solar], while the heavier elements had almost normal abundances. Figure 7-17 (Antipova

1974) shows the abundances of DQ Her compared with solar ones.

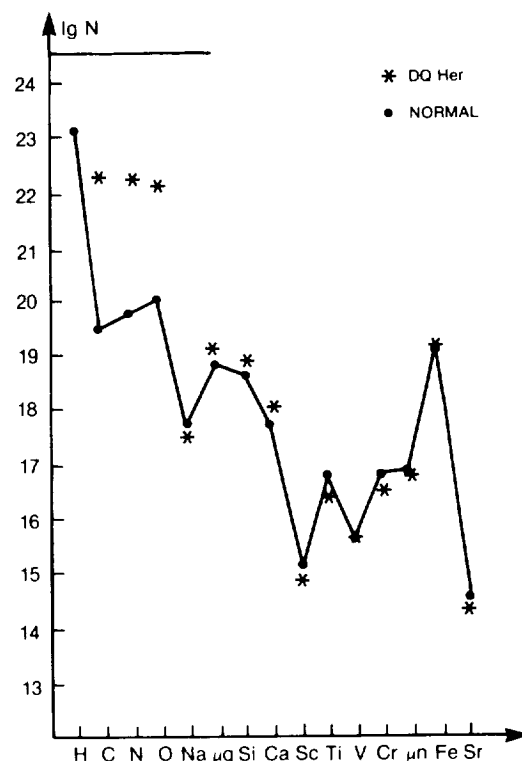


Figure 7-17. The chemical composition of the envelope of DQ Her, derived from outburst spectra, and compared with solar values (Antipova 1974).

It is rather dangerous to apply classical curve of growth theory in a situation where non-LTE and nonthermal effects can be expected. The similarity of the nova spectrum with that of an F supergiant may be deceptive, as can be seen when the above discussion of possible models is kept in mind. The more detailed examination of Williams (1977) leads to other criticisms. The CNO abundances were derived from very few lines (at most, 5 per element); indeed, there were only enough lines to construct separate curves of growth for the ions Fe I, Fe II, Ti II and Cr II. Mustel and his co-workers assumed the same curve of growth for all ions with the same "microturbulent" velocity, and the same law of atom/ion population variation with excitation potential. They had to assume that the latter deviated from a Boltzmann distribution, T_{exc} increasing with excitation potential. Such assumptions can lead to very uncertain abundances, especially

when derived for H, C, N, and O, with lines having a high excitation potential. Though the measurements appear to be consistent with the mentioned assumptions and also with the use of the Saha equation, results derived from the curve of growth method should be viewed with much caution.

The absorption spectrum of DQ Her has also been studied in another way by Sneden and Lambert (1975). This nova showed CN absorptions during early postmaximum development, and spectral synthesis could be used to obtain information about isotopic ratios. The CN lines of the first six vibrational bands of the $\Delta v=1$ sequence of this molecule were analyzed, assuming formation in a scattering layer (fractional transmission proportional to $e^{-\tau}$ with τ being the optical depth). Thermal equilibrium was assumed for the molecule, with an arbitrary microturbulence of 5 km s^{-1} , atomic lines being neglected. The isotopic ratios obtained were $^{12}\text{C}/^{13}\text{C} \geq 1.5$ and $^{14}\text{N}/^{15}\text{N} \geq 2$, which can be compared with the corresponding solar system values of 89 and 273.

IV.D. ELEMENT ABUNDANCES IN NOVA EJECTA: NEBULAR LINES

In the post maximum stage, the spectrum of a nova exhibits a large number of emission lines, which in principle, can, be used to derive physical and chemical properties of the ejected envelope (or merely, the ejected shells that may be interacting), as well as properties of the central object. In general, lines originating from transitions requiring lower densities and higher levels of ionization appear at later stages of the development.

When the nova enters the nebular stage, it is generally not spatially resolved. Integrated intensities of different emission lines, which refer to the entire envelope as if it were a homogeneous nebula, are used for the analysis. This assumption is certainly not correct and leads to the most serious objection to an uncritical application of plasma diagnostics. A remedy would be a high resolution spectrophotometric study of emission lines, resolving at least the structure in the line of sight. Previously such

observations have been very scarce.

The flux emitted by any line depends on how it is excited; knowing the physical conditions, one can deduce the abundance of the ion to which it belongs. Basically one uses the following equation (Collin-Souffrin, 1977):

$$I_{\lambda} = \frac{1}{4\pi d^2} \int f_{\lambda}(T_e, n_e) n(X_{ij}) dV, \quad (7.16)$$

where I_{λ} is the absolute flux of the line ($\text{ergs cm}^{-2} \text{ s}^{-1}$) at the earth, d is the distance of the nova, V is the volume of the envelope, $n(X_{ij})$ is the number density of the ion X_i in the j^{th} state of excitation giving rise to the line, and f_{λ} is the emissivity of the line at an electron temperature T_e and an electron density n_e . This equation needs to be corrected for reddening and line self-absorption when present. It is clear that the observed inhomogeneities can pose very serious problems when trying to apply Equation (7.16). In addition one needs to know the abundances of all ions of an element in order to determine the total abundance; states of ionization not giving rise to lines in observed spectral regions are difficult to include in calculations, particularly when there are uncertainties concerning the ionization. The values of T_e and n_e are derived from line ratios that are sensitive to either of the parameters, e.g., $([\text{O III}] 4959, 5007)/([\text{O III}] 4363, [\text{O III}] 4959, 5007)/[\text{Ne III}] 3869$, $([\text{O III}] 4959, 5007)/(\text{He I } 5876)$. The ionization fractions are assumed to be time-invariant (Seaton, 1975; Ferland and Shields 1978b; Lance et al., 1988). The relevant atomic constants can be taken from a compilation by Mendoza (1983).

Studies of abundances, like studies of the physical conditions in general of regions of emission line formation, are best based on results from many different parts of the spectrum. Combinations of satellite ultraviolet observations with optical ones are better than those based only upon the latter.

Filling factors in nova shells appear to be low, since most of the high-density material responsible for the emission is in clumps, filaments, or thin sheets, while the material in

between is at much lower densities and presumably higher temperatures. Masses in the past might have been overestimated by a factor of 10 (Peimbert and Sarmiento, 1984). However, the in-between material would be hard to detect, and its mass could be underestimated.

IV.E. INTERPRETATION OF INFRARED OBSERVATIONS

Infrared observations in different stages yield the following results:

Around optical maximum, the infrared flux of a nova yields the Rayleigh-Jeans tail of the pseudophotosphere, usually at a temperature of 7,000 - 10,000 K. In early decline, it transforms into an optically thin free-free (thermal Bremsstrahlung) flux distribution. At this phase, emission lines (or emission bands) often are superimposed on the continuum (see below for details). In many cases, this free-free emission is replaced by blackbody radiation from a circumstellar dust cloud of a temperature of typically 1,000 K. It condenses from the nova ejecta, or (as an alternative explanation, which encounters more difficulties), is a cloud of preexisting dust, heated by the radiation of the outburst.

Soon after the peculiar drop in the light curve of DQ Her, McLaughlin (1935, 1937) suggested that a cloud of dust had suddenly formed from the ejecta, producing the dramatic fading of the visual flux. However, since no infrared observations were made at that time, the suppression of the redshifted components of the emission lines, which is also today considered as a good criterium for dust formation, was then the only evidence for the existence of such a dust cloud.

Only in 1970, with the thorough study of the energy distribution of FH Ser in different stages of the outburst did the explanation of light curve disturbances by dust won acceptance: a fading of optical and ultraviolet flux coincided with the emergence of strong infrared continuum emission, balancing the energy output completely.

Bode and Evans (1983) sorted the novae according to their infrared development into three classes:

Class X: novae for which, at a certain stage, the infrared luminosity is nearly the luminosity of the underlying object. These novae invariably have a pronounced discontinuity in the visual light curve, which coincides with the onset of infrared development. The temperature of the dust shell attains a minimum before rising to a plateau—the so-called isothermal phase (example: NQ Vul).

Class Y: novae for which the infrared luminosity is below 10% that of the underlying object. The light curve is smooth and the dust shell temperature decreases monotonically (example: V1668 Cyg).

Class Z: novae with little or no infrared excess, i.e., little or no dust. The visual light curve is usually smooth. The excess can easily be attributed to line emission in the infrared (example: V1500 Cyg).

As has been shown by Bode and Evans (1981), the onset of infrared excess occurs (if it occurs) usually when the nova has declined by 4 magnitudes, and shortly after the transition phase of spectral development.

IV.E.1 GRAIN GROWTH

In the model of Clayton and Wickramasinghe (1976), grains can only condense from ejecta once the temperature of condensation is such that the saturated vapour pressure of the grain material is less than the partial pressure of the ambient monomer gas. This condition is fulfilled at 2,000 K, and grains begin to grow. Their temperature declines for two reasons: first, the distance to the central object is increasing, and second, the absorption efficiency of the grains increases as grains grow. Growth of an individual grain is essentially complete in a few days, and maximum grain size is 2 μm .

Because the photospheric temperature of novae increases after outburst, the distance at

which grains can condense recedes from the nova. Dust formation thus depends on the question whether the ejecta, which are obviously also receding, can overtake the condensation distance. If so, grain formation is possible; otherwise, the nova environment is always too hot. Furthermore, as the ejecta disperse, their density declines, and unless grain growth is initiated at an early stage, the amount of dust eventually formed is negligible.

Gallagher (1977) developed an idealistic model by assuming that the condensation of grains is controlled by the radiation field of the nova, which radiates at constant luminosity. Its surface temperature can be estimated to be

$$T_s = 5100 (L_{CL}/L_\odot)^{1/4} \text{ K}, \quad (7.17)$$

where L_{CL} is the luminosity at outburst. The temperature of a grain can be written as

$$T_g^4 = \frac{1}{4\sigma} \frac{L_{CL}}{4\pi R^2} \frac{Q(T_g, a)}{Q(T_s, a)}, \quad (7.18)$$

where Q is the Planck mean absorption cross section for grains of radius a and temperature T_g in a radiation field characterized by the stellar photospheric temperature T_s . Assuming that the early grains are simple,

$$T_g \approx (L_{CL}/4\pi\sigma r^2)^{1/4}, \quad (7.19)$$

This can be written in terms of the time t_d at which dust initially becomes observable:

$$t_d \approx \frac{2}{T_g^2(0)} \sqrt{\frac{L_{CL}}{4\pi\sigma}} \frac{1}{V}, \quad (7.20)$$

where V is the expansion velocity.

Calibrating this formula with FH Ser, adopting $T_g(0) = 1300 \text{ K}$, this yields:

$$t_d = (320/V) \sqrt{(L_{CL}/L_\odot)} \text{ days}, \quad (7.21)$$

Fast novae have several properties that limit grain formation. First, the ejection velocities are larger than in slower novae, producing lower gas densities at a given time (especially

since there is no evidence for the ejection of larger masses in fast novae). Second, fast novae produce substantial stellar winds, i.e., high-velocity material, that may give rise to shocks that might disrupt the grain nucleation process. Third, ionization might be higher.

The ionization time scale is given by Gallagher to be

$$t_i = 2.2 \cdot 10^6 / V (L_{CL}/L_\odot)^{1/3} \text{ days}. \quad (7.22)$$

The ratio of t_i/t_d is independent of the expansion velocity and depends only on the luminosity or speed class of novae. For fast and very fast novae, $t_i < t_d$, and dust does not form. In moderately slow and slow novae, the equation for t_d gives a reasonable estimate for the time of dust formation. In this picture, HR Del, as well as most other slow novae, are not understood, because they form too little dust for their slow speed.

IV.E.2. LINE OR BAND EMISSION

The most spectacular line in the infrared observed until now is the $12.8 \mu\text{m}$ [Ne II] emission in QU Vul, originating from the transition $P_{1/2} - P_{3/2}$, which amounted to 0.1% of the outburst luminosity at a given date (Gehrz et al., 1985, 1986). From Ne III and Ne IV lines in the ultraviolet simultaneously present, a high overabundance of Ne could be derived. Together with SiO features observed at $10 \mu\text{m}$ and the suspicion that an overabundance of Mg is hidden in the 10 and $20 \mu\text{m}$ features, QU Vul is a good candidate for a TNR on an O-Ne-Mg white dwarf (see Chapter 7, Section V.A.) with a large amount of white dwarf material mixed into the ejected shell.

Less spectacular features are the $5 \mu\text{m}$ emission, a short-lived feature in the early free-free phase, which was observed in V1668 Cyg and LW Ser. It is attributed to CO (Ferland et al., 1979). A fairly mysterious broad emission feature around $10 \mu\text{m}$ occurred in V1301 Aql, which is attributed to SiC or possibly CS.

IV.F. INTERPRETATION OF RADIO OBSERVATIONS

Radio emission from novae (HR Del and FH Ser) was first discovered in 1970 (Hjellming and Wade, 1970).

The radio light curve for novae develops much more slowly than the optical one. Let us assume a thermal absorption coefficient κ_ν , which should take into account the usual chemical composition of the nova shell (e.g., Scheuer, 1960).

The radial optical depth is given by

$$\tau_\nu = \int \kappa_\nu dr. \quad (7.23)$$

At a fixed date t , $\tau_\nu(t)$ behaves like ν^2 , and at a fixed ν , it behaves like t^5 , assuming expansion in the form of a Hubble flow (instantaneous ejection II).

For ν and t sufficiently small, i.e., at radio wavelengths, we have $T_\nu(t) \gg 1$, and the nova radiates like a blackbody:

$$S_\nu = \pi \left(\frac{R}{r} \right)^2 B_\nu(T_e) \quad (7.24)$$

(thermal emission for $\tau_\nu \gg 1$),

where R is the outer radius of the shell and r , the distance. For ν and t sufficiently large, $\tau_\nu \ll 1$, and

$$S_\nu = \int_{\text{volume}} \kappa_\nu B_\nu \frac{dV}{r^2} \quad (7.25)$$

(thermal emission for $\tau_\nu \ll 1$)

Thus, for Hubble flow expansion, at this stage, $S_\nu \sim D^{-3}$.

Hjellming et al. (1979) have calculated the radio development of an outburst and compared models with observed radio light curves. From this comparison, they derived information regarding mass, velocity, and radio thickness of the shell of V1500 Cyg. The assumed model is spherically symmetrical, with sudden ejection of an isothermal ($T_e = 10,000$ K) shell

with a velocity gradient. For V1500 Cyg, the model gives a mass of $2.4 \cdot 10^{-4} M_\odot$, the velocities are 200 and 5600 km s⁻¹ at the inner and outer radii, respectively. The shell becomes radio thin after some 100 days. Infrared fluxes and their evolution in time are also well predicted.

Hjellming et al.'s models of slower novae give lower velocities of ejecta, lower masses of the shell, and greater time elapsed from outburst for the shell to become radio thin.

IV.G. STRUCTURE OF NOVA SHELLS

For a few nearby old novae, shells have been observed in quite some detail, and deviations from spherical symmetry are always encountered. Nonspherical envelopes were explained by

- Magnetic guiding of material (Mustel and Boyarchuk, 1970).
- Nonradial stellar pulsations, which is still a very debatable assumption (Warner, 1972).
- Interaction of the expanding nebula with the secondary component: blobs ejected perpendicular to the orbital plane (Hutchings, 1972; Pilyugin, 1986).
- Interaction of the expanding nebula with the accretion disk (Gorbatskii, 1974; Sparks and Starrfield, 1973).
- Effects of stellar rotation, gravitational braking, and radiative acceleration (Phillips and Reay, 1977)
- Rayleigh-Taylor instabilities at the beginning of the expansion (Chevalier and Klein, 1978)
- Stellar rotation only; TNR proceeds at the same rate at all latitudes; radial ejection velocity is the same at all latitudes. The oblateness of shells of about 1.5 can be explained with a rotating white dwarf with

a period of the order of minutes. No radiative acceleration is needed; it could also be shown that the accretion disk, hardly, and the secondary, only to a small extent, influence the kinematics of the (principal) ejection. As the shell expands, it must initially cool rapidly and may be subject to thermal instabilities. This tends to form concentrations of characteristic size about $(c/v_{\text{eject}})r$. Since the Mach number of the ejecta is probably initially about 1, the shell should contain only a few condensations (Fiedler and Jones, 1980).

V. CAUSES OF NOVA OUTBURSTS

The theory at present accepted by almost all workers in the field is one where hydrogen is accreted by the white dwarf component of the binary from its companion, and where this hydrogen undergoes a thermonuclear runaway. The detailed description of the theory is beyond the scope of this book devoted to atmospheres; what will be emphasized is the impact of it on observable properties. More details will be found in the reviews by Sparks et al. (1977), Starrfield and Sparks (1987), and Starrfield (1988).

V.A. MODELS FOR CLASSICAL NOVAE

Hydrogen accreted by a white dwarf will tend to be ignited, and as this process accelerates, a thermonuclear runaway can eventually occur. This is possible because there should be no significant transport of energy into the interior of the white dwarf from the outer regions where hydrogen is burning. The strength of the outburst is determined by a proper pressure at the core-envelope interface of the white dwarf

$$p = \frac{GM}{R^2} - \frac{\Delta M}{4\pi R^2}, \quad (7.26)$$

the right-hand side being the gravitational attraction multiplied by the mass of the accreted material per unit area. In this expression, ΔM is the envelope mass, M is that of the white

dwarf, and R its radius. MacDonald (1983) and Fujimoto (1982) found that a value of p of 10^{20} dynes cm^{-2} is necessary for a fast nova outburst. Using a white dwarf mass-radius relation, one deduces a relation between the envelope mass and the white dwarf mass required for a fast nova outburst; the envelope mass decreases as the white dwarf mass increases. It becomes easier to produce a nova outburst when the white dwarf is more massive.

In reality, the situation is less simple. The mass accretion rate, the chemical composition, and the white dwarf luminosity all influence the evolution of what will give rise to an outburst. The influence of the mass accretion rate on the development of accreting white dwarfs having a solar composition will be discussed in chapter 12, where possible models for symbiotic stars are considered; the theory is not relevant only for classical novae. When the accretion rate of the white dwarf component increases, the mass of the accreted envelope decreases (cf., MacDonald, 1980) because of the gravitational compression of the accreted material, which produces energy that accelerates the thermonuclear runaway. It appears, moreover, that classical nova explosion are not produced for high accretion rates above $10^{-8} M_{\odot} \text{ yr}^{-1}$ unless the white dwarf mass is very close to the Chandrasekhar limit; in view of the observational evidence for such accretion rates, this poses a problem we will discuss in more detail.

While variation in the mass accretion rate is associated with variation in the energy release due to gravitational compression of accreted material by white dwarf components, the abundances of CNO are also very important for the physics of nova explosions. This can be understood theoretically because of the influence of the β unstable nuclei ^{13}N , ^{14}O , ^{15}O , ^{17}F . At first their lifetimes (863, 102, 176, and 925 s) are shorter than the CNO nucleus lifetime against proton capture, so the β unstable nuclei can quickly decay without holding up the following nuclear reactions. At later stages in the development of a prenova, the temperature rises, the proton captures that precede and succeed β decays are more rapid, and a β decay bottleneck

can build up. The energy generation is proportional to the CNO initial abundance at this point, where the temperature is approximately 10^8 K, as new CNO nuclei are not created but only redistributed. We should also note that the time delay between the creation of the β unstable nuclei and their decay leads to the storage of energy, so energy can be released after the envelope has begun to expand ($10^2 - 10^3$ s after production of the β unstable nuclei). In fact, the calculations suggest that an overabundance of CNO is necessary for a fast nova outburst, i.e., mixing of CNO into the envelope from the interior occurs. This is perhaps the place where theory and observation of novae come closest, and where many would begin at least to suspect that they are not talking of completely different things.

According to present theory, the initial luminosity of the white dwarf has no influence on the amount of accreted mass, when it is low. Energy is then generated from the proton-proton chain for which secular evolution of the envelope is very slow. When the time scale for nuclear burning is larger than that for accretion, the rate of evolution of a prenova is determined by the rate of mass accretion. However, for a higher initial luminosity, nuclear burning is from CNO reactions and their time scale can be shorter, thus influencing the mass accreted before a thermonuclear runaway.

For a thermonuclear runaway to produce ejection, the material of the shell source must be electron degenerate. The kinetic temperature of the gas must rise and exceed the Fermi temperature before expansion can occur and cool the gas. The Fermi temperature is given by Starrfield et al. (1985) as equal to

$$T_F = 3 \times 10^7 \left(\frac{\rho_i}{\mu_i} \right)^{2/3}, \quad (7.27)$$

with ρ_i the density in units of 10^3 g cm $^{-3}$ and μ_i the mean molecular weight multiplied by the ratio of the number of all particles to that of electrons. If the temperature is rising rapidly, expansion only begins to stop the temperature rise when the temperature is much larger than

T_F . The calculations indicate that convection is present during evolution to the peak of the outburst. The convection turnover time scale is of the order of 10^2 s, and β unstable nuclei reach the surface before decaying. Fresh unburnt material is brought into the hot shell source by the convection, and the β unstable nuclei are the most abundant of the CNO nuclei at outburst peak.

Very many calculations of models have been carried out and will not be described here. Detailed calculations, for instance were carried out by Starrfield et al. (1978), Sparks et al. (1979), and by Starrfield et al. (1985, 1986). Realistic velocities and ejected masses were obtained; however, this is not sufficient to demonstrate the validity of these models. Their description of postoptical maximum development will be perhaps a more sensitive test of future models.

Up to now, calculations suggest an expansion of the white dwarf component after the initial explosion. This component should engulf its companion and radiate at a luminosity not far from the Eddington limit for quite a long time. Therefore, this type of theory appears to be most compatible with central star-dominant models describing postoptical maximum evolution. Such a central star could have a strong wind like more "normal" hot stars; however, as already described, novae generally appear not to be like this. Another effect can be expected if the secondary revolves inside an extended white dwarf. Gravitational stirring would take place with extra energy generation, and perhaps the total luminosity could then exceed the Eddington limit. This last situation has been discussed by MacDonald (1980) and MacDonald et al. (1985), but such calculations are still too simple. In particular, deviations from spherical symmetry need to be properly taken into account. It is in the framework of such considerations that there is a possibility of understanding the production of an optically thick wind at large optical depths. Future work may or may not confirm this.

Various developments of thermonuclear

runaway theory have occurred in the 1980s. The observed apparent overabundances of Ne, Mg, etc., for some novae have lead to a certain amount of work on this second class of nova outburst. Delbourgo-Salvador et al. (1985) and Starrfield et al. (1986) explain them by nova outbursts of O-Ne-Mg white dwarfs. Such white dwarfs are expected to have masses close to the Chandrasekhar limit, the lower limit to their masses being somewhat uncertain. Only a small proportion of white dwarf components (on the order of a few percent) should be of the O-Ne-Mg class. The calculations indicate, however, that nova outbursts would be more frequent, so as many as 20% of observed outbursts could be of this type. Material of the white dwarf, as for other types of novae, would be mixed into the envelope. These stars are expected to be more massive, leading to runaways with less accreted mass and so occurring more frequently.

V.B. NONSPHERICAL MODELS

Accretion is not spherically symmetric, and this lack of spherical symmetry should play a role in the development of a prenova. Kippenhahn and Thomas (1978) considered the formation of a rapidly rotating belt following accretion; its chemical composition and angular momentum are then mixed with the underlying white dwarf material. Marginal stability, with respect to the Richardson criterion for shear instability, was assumed to be maintained by mixing, the amount of mixing at a certain time depending on the position and depth inside the belt. Hydrogen would be eventually ignited at the bottom of the belt according to this scenario. Kutter and Sparks (1987) and Sparks and Kutter (1987) considered accretion of material possessing angular momentum with fewer assumptions than Kippenhahn and Thomas, but still including marginal instability against shear mixing. Various cases between radial accretion and material accreted having a full Keplerian orbital velocity were considered. Unfortunately, a $1 M_{\odot}$ white dwarf did not produce nova-like mass ejection under these conditions. The authors explain this lack of success, compared with other types of model, to

the support the centrifugal force gave to the accreted matter, diminishing pressure confinement and the strength of the runaway. Kutter and Sparks considered that other physical effects needed to be taken into account in later work.

Shaviv and Starrfield (1987) considered another aspect of deviation from spherical symmetry. The boundary layer between the accretion disk and the white dwarf can (and was also supposed to) cause heating—a nuclear burning region in the whole accreted envelope was produced, and the latter become completely convective. Unfortunately, “no dynamical effect occurred during the evolution.” By adding another badly known parameter, the degree of heat flow inward from the boundary layer, the authors admit to “have complicated an already cloudy situation.” It is clear that much work needs to be done.

V.C. OTHER EFFECTS AND HIBERNATION

Another feature of some models of the mid-1980s for classical novae also needs to be discussed. Some observational and statistical evidence suggests that novae “hibernate” during outbursts, i.e., that very old novae become very faint for millenia and brighten again before the following thermonuclear runaway. Such ideas were suggested by the accretion rates in old novae, deduced from their energy distribution, which appear to be too high for nova explosions, according to the simple model previously discussed. In addition very old novae appear to be fainter than more recent novae, while, at one time, the lack of X-ray sources expected for a large population of old novae was considered to be a problem.

Hibernation models have been reviewed by Livio (1987). It is supposed that, following a nova explosion, the separation of the binary increases. At first, the secondary continues to transfer mass because it is strongly irradiated. This mass transfer then strongly decreases because the secondary underfills its Roche lobe, allowing previously accreted material to

cool, diffuse, and become degenerate so that a strong thermonuclear runaway is possible at a later time. The separation of the binary however, is reduced by magnetic braking (if the period is above the cataclysmic binary period gap) and returns to its original value and a high mass accretion rate on a time scale

(7.28)

$$\tau = 5000 \left(\frac{f}{0.7} \right)^2 \left(\frac{r_g}{0.45} \right) \left(\frac{R_{wd}}{6 \times 10^8 \text{ cm}} \right)^4 M^{-3/3} P_4^{+0.96} \text{ yr.}$$

Here f is a parameter of the order of 0.7, and r_g is the gyration radius of the secondary. R_{wd} is the radius of the white dwarf in cm, M is the total mass of the system in M_\odot , and P_4 is the period in units of 4 hours. It is after such a time that the white dwarf resumes accretion to produce a new thermonuclear runaway. At stages when the mass transfer is still relatively low, dwarf nova outbursts should be possible if disk instability models are a valid explanation for them, so classical and dwarf novae would be the same objects. Livio (1987) indeed lists eight old classical novae with post outburst eruptions similar to those of dwarf novae.

According to Shara et al. (1986), the separation increase of the binary is due to the mass ejection in the nova explosion, and this effect is larger than that due to angular momentum loss produced by interaction of the revolving secondary with the mass ejected by the nova. This was supposed by the period increase observed for one nova (BT Mon); suitable data do not exist for other novae. Such a discussion clearly neglects the possibility of the white dwarf expansion and the engulfment of its companion.

The present situation concerning the relevance of hibernation models is not clear. Recent results may indicate a much smaller effect on the mass transfer than previously thought. If the atmosphere of the companion is isothermal due to irradiation by the white dwarf, a mass transfer decrease by a factor of the order of 10 - 100 can be expected, but if the atmosphere is convective, the decrease is at most by a factor of 2. In the former case of "mild hibernation," the mechanism still works (Livio 1988a). If

there is no hibernation, it may be necessary to suppose the accretion rates deduced from the luminosities of old novae wrong; the white dwarf could be exceptionally bright before and after the explosion, and its radiation could be reprocessed by the accretion disk (Friedjung 1985, Livio 1988b).

If we consider thermonuclear runaway theories in general, some success has been achieved. As predicted, fast novae have overabundances in CNO, the overabundance being correlated with the speed of development of a nova. Even the CNO overabundances of the slow nova DQ Her can be explained by the exceptionally low mass of the white dwarf. Postmaximum activity is probably also compatible with such models, but detailed predictions do not exist. However, when one attempts to take account of other physical effects, difficulties are encountered. A theory that explains most observations is still far away. In any case, theoretical prejudices should not be used as a pretext for rejecting models based on observations.

Unlike with other types of variable stars, the determination of the space density of novae is extremely difficult to determine, since year after year, new nova explosions are discovered. Thus, a straightforward density of observed novae that showed outbursts would be an ever increasing function of time. A parameter that can more easily be determined is the space/time density ρ^* , that is the number of nova explosions per cubic parsec per year. A determination of this parameter, as well as the scale height of the nova distribution in the Galaxy, is given by Duerbeck (1984).

To derive the true space density ρ , the mean time interval $\Delta\tau$ between two nova outbursts must be known. This value is extremely uncertain, and the only safe statement is that it is obviously larger than 100 years for classical novae.

If we assume that the nova state is a steady one, with the matter ejected during outburst being equal to the mass accreted between out-

bursts, values of $\Delta\tau$ between 1,000 and 10,000 years are derived for the observed shell masses and accretion rates. However, Prialnik (1986) and Kato (1988) have argued from calculations of nova outbursts that a secular loss of matter from the white dwarf that undergoes nova explosions takes place, thus shortening the interval between explosions. On the other hand, theoretical arguments for orbit changes during outbursts, the impossibility of TNRs under highly degenerate conditions at the accretion rates observed, and observational findings from the (unfortunately few) very old novae lead to the suspicion that mass transfer rates may diminish noticeably or may even cease. Thus, for a shorter or longer time between outbursts, the exnova transforms into a cataclysmic binary with possible disk instabilities (i.e., a dwarf nova), or even in a detached system, consisting only of the red dwarf and the white dwarf, with the accretion disk completely absent. Such systems would appear as red dwarfs with UV excess (possibly eclipsing) and rapid rotation, properties that are not obvious in low-dispersion spectral surveys. Such a decrease in mass transfer rate would increase the outburst interval.

Taking a mean outburst interval (without hibernation) to be 3,000 years, a space density of

$$\rho_0 = 1.27 \cdot 10^{-6} \text{ pc}^{-3} \quad (7.29)$$

is derived for classical novae, which can be compared to that of dwarf novae, $0.95 \times 10^{-6} \text{ pc}^{-3}$, and symbiotic stars, $0.00094 \times 10^{-6} \text{ pc}^{-3}$, the scale height of the latter however, being very different from the first two groups.

If these assumptions are correct, one should find, if the nova brightness between outbursts stays at $V = +4.2$, 6×10^{-2} objects in the bright star catalogue (brighter 6^m), 5 - 6 objects in the Durchmusterung catalogues (brighter 9.5^m), and 36 objects brighter than 11^m . While there are some 'novalikes' in the brightness range of $9^m - 11^m$ their number and type makes them not too well suited for nova candidates.

However, the hibernation scenario is not a

good scenario out of the dilemma. Decreasing the accretion rate means decreasing the absolute magnitude, thus decreasing the volume of space where the objects are found. On the other hand, a lower accretion rate means that the intervals become longer, and the space density of the objects must be increased to result in the same—observed—outburst density ρ^* .

V.D. CAUSES OF RECURRENT NOVA OUTBURSTS

Recurrent novae to a certain extent are intermediate between classical and dwarf novae. Indeed, there has sometimes been a certain amount of confusion about whether a particular star is a recurrent nova or a dwarf nova. To clearly make the distinction, Webbink et al. (1987) consider that a recurrent nova has two or more recorded outbursts with a maximum absolute magnitude comparable to that of a classical nova ($M_v \leq -5.5$) and ejection of a discrete shell in outburst with an expansion velocity of $\geq 300 \text{ km s}^{-1}$. These criteria also distinguish recurrent novae from various types of symbiotic stars. Even with such criteria, recurrent novae seem to be rather heterogeneous; the outbursts of some are now explained by theories similar to those for classical novae, while the outbursts of others are explained by theories bearing some resemblance to those of dwarf novae.

If we try to invoke a mechanism similar to that for classical novae, i.e., involving a thermonuclear runaway, rather stringent conditions need to be fulfilled. One can see from equation (7.26) that the pressure at the core interface is inversely proportional to the fourth power of the radius of the white dwarf; this will increase very rapidly near the Chandrasekhar limit, where the white dwarf radius becomes very small. Hence, a runaway will be produced for a relatively small amount of accreted mass, so if the accretion rate is supposed not to change very much, much shorter recurrence times are possible near the Chandrasekhar limit. In this way, short recurrence times are possible without the accretion rate becoming too high, to produce a strong outburst.

An increase of the white dwarf luminosity also leads to shorter recurrence times. Starrfield et al. (1985) for a limiting mass of $1.38 M_{\odot}$, a luminosity of $0.1 L_{\odot}$ and an accretion rate of $1.7 \times 10^{-8} M_{\odot} \text{ yr}^{-1}$, were able, to obtain a recurrence time of only 33 years. A rather high accretion rate would lead to a high accretion disk luminosity, detectable between outbursts, supposing naturally that accretion proceeds via a disk, Webbink et al. (1987) give for a $1.38 M_{\odot}$ white dwarf:

$$\frac{L_{\text{bol}}}{L_{\odot}} \frac{GM_{\text{WD}}}{R_{\text{WD}}} \approx \frac{160(M_{\text{WD}}/1.38 M_{\text{WD}})}{R_{\text{WD}}/1.9 \cdot 10^8 \text{ cm}} \frac{\dot{M}}{10^{-8} M_{\odot} / \text{yr}}, \quad (7.30)$$

with \dot{M} the accretion rate, L the stellar luminosity, M_{WD} the stellar mass and R_{WD} the stellar radius.

As in the case of dwarf novae, recurrent novae, in principle, can, also be produced by accretion events. Such events might be powered by an instability of the cool component, or by a disk instability, or they might occur at periastron if the companion had an eccentric orbit. In the last case, the eccentricity must exceed the ratio of the pressure scale height near the inner Lagrangian point at periastron to the radius of the star losing mass by Roche lobe overflow. It can be noted that when accretion events occur, unlike in a thermonuclear runaway, the accretor need not be a white dwarf with a mass below the Chandrasekhar limit. It can be a main sequence star, as is indeed indicated by the most probable compact star mass above the Chandrasekhar limit for T CrB (see Chapter 9 of this volume).

Different accretion event mechanisms for recurrent novae can be examined in more detail as was done by Livio (1988). According to him, the accretion rate \dot{M} must obey the condition for disk instability to occur:

$$\dot{M} \geq 3 \cdot 10^{-9} \left(\frac{P_4}{4} \right)^{1.8} M_{\odot} \text{ yr}^{-1}, \quad (7.31)$$

with P_4 the period in units of 4 hours. So for a period of 230 days (T CrB, RS Oph), $\dot{M} \leq 10^{-6}$

$M_{\odot} \text{ yr}^{-1}$. Disk instability models, however, may run into recurrence time problems; to obtain times of the order of those observed, a viscosity parameter α of the order of 10^{-3} is required (for the cold state of the disk).

A more relevant accretion event mechanism according to Webbink et al. (1987) and Livio (1988a) for some recurrent novae involves a sudden instability of the cool component with ejection of 10^{-3} to $10^{-4} M_{\odot}$. This mechanism is supposed to be particularly relevant for T CrB and RS Oph, and according to Edwards and Pringle (1987) to be possible for Roche lobe filling giants. The collision of the ejected material with itself, with the cool giant wind, or with the accreting star can be associated with high-velocity shock ejection, when no well-developed accretion disk exists before the event. According to Livio et al. (1986), the shock velocity is given by

$$V_{\text{shock}} = \left(\frac{2 GM}{R} \right) \left(\frac{\rho_{\text{circum}}}{\rho_{\text{stream}}} \right)^{\frac{1}{2} \left(\frac{2\gamma}{\gamma+1} \right)^{0.5}} \quad (7.32)$$

The first factor, giving the free fall velocity from zero at distance R from a star of mass M , is multiplied by the second factor, including the ratio of the density of circumstellar material ρ_{circum} to that of the stream of accreted material ρ_{stream} , which collides with it. As usually, γ is the ratio of specific heats, and G , the gravitational constant. Therefore, the highest shock velocity occurs if the stream collides with very low-density material in the vicinity of the accreting star. The accreted material can then form a temporary bright disk (whose absence before outburst would pose a problem for thermonuclear runaway models), explaining the secondary maximum of T CrB, or it can easily collide with the accreting star if it is bloated, following considerable accretion at a high rate (suggested for RS Oph). It should be noted that 10^{-4} of the accreted material (according to the low envelope mass estimate of Bode and Kahn (1985) for RS Oph) needs to be ejected at a velocity of 10 times the free fall ve-

locity of RS Oph (of the model of Livio et al., 1986, with R equal to 8.2×10^{11} cm), so then requiring only a high efficiency for the conversion of the kinetic energy of the accreted material of the order of 10^{-2} . This type of mechanism clearly needs to be studied in more detail.

Webbink et al. (1987) support thermonuclear runaway models for T Pyx and U Sco. The former, to some extent, resembles a classical slow nova. The presence of bright accretion disks, characteristic as an accretion rate preceding a thermonuclear runaway, is compatible with observations for both these recurrent novae, though Webbink et al., suggest that most of the quiescent luminosity of T Pyx has another source (continuing nuclear burning?).

In a later paper, however, Truran et al. (1988) are not so certain about U Sco; observations appear to indicate a high He/H ratio for which it is difficult to produce a nova-like outburst following a thermonuclear runaway. These authors suggest that a high helium abundance could favor an accretion disk instability even for a hot accretion disk.

One can conclude that models for recurrent novae need to be compared with results for individual stars. The situation for recurrent novae is not clear; one may wonder whether other possible mechanisms for outbursts have not been neglected. One tends to gather the impression that at a give time, theorists are too certain about their mechanisms.